AD A118346

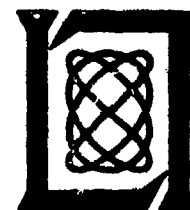| Technical Report	582
Application of Advanced Signal Processing Techniques to Angle of Arrival Estimation in ATC Navigation and Surveillance Systems	J.E. Evans J.R. Johnson D.F. Sun 23 June 1982

DTIC
ELECTE
AUG 1 7 1982
S
H

82 08 16 249

1. Report No. FAA-RD-82-42	2. Government Accession No. AD-A118 306	3. Recipient's Catalog No.	
4. Title and Subtitle Application of Advanced Signal Processing Techniques to Angle of Arrival Estimation in ATC Navigation and Surveillance Systems		5. Report Date 23 June 1982	
		6. Performing Organization Code	
7. Author(s) J.E. Evans J.R. Johnson D.F. Sun		8. Performing Organization Report No. Technical Report 582	
9. Performing Organization Name and Address Massachusetts Institute of Technology Lincoln Laboratory P.O. Box 73 Lexington, MA 02173-0073		10. Work Unit No.	
		11. Contract or Grant No. DOT-FA74-WAI-461	
12. Sponsoring Agency Name and Address Department of Transportation Federal Aviation Administration Systems Research and Development Service Washington, D.C. 20591		13. Type of Report and Period Covered Technical Report	
		14. Sponsoring Agency Code	

15. Supplementary Notes

16. Abstract

This report focuses on the mitigation of multipath and/or incoherent (cochannel) interference on receiving antenna arrays since these phenomenon represent a principal challenge to many existing ATC systems. The improved signal processing techniques considered explicitly assume that interfering signals may be present and estimate the interference parameters (e.g., power and angle of arrival) as well as the desired signal parameters. By thus identifying the nature of the interference, it is then possible to reduce its effects on the desired signal angle of arrival estimates.

The specific topics addressed include:

(1) analytical calculations of the theoretically achievable performance as a function of the number of elements in the array, number of interfering signals, and signal-to-noise ratio (SNR),

(2) assessment of the comparative performance of representative techniques, including those suggested by the Federal Republic of Germany for the DME-Based Landing System (DLS) as well as representative US techniques (maximum likelihood, autoregressive modelling (AR), eigenanalysis), via analytical studies, statistical decision theoretic computer simulations, and experimental measurements,

(3) hardware implementation features such as equipment complexity, calibration, and monitoring, and

(4) investigation of how promising techniques might be applied to the angle guidance processing for the Microwave Landing System (MLS).

The principal conclusions can be summarized as follows:

(a) There seems little likelihood of significantly improved angle estimation accuracy for cases where only the desired signal is present and signal to noise ratio is the principal limitation.

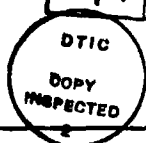| 17. Key Words High Resolution Spectrum Estimation Microwave Landing System (MLS) Multipath Maximum Entropy Autoregressive Modeling Maximum Likelihood Eigenanalysis	18. Distribution Statement Document is available to the public through the National Technical Information Service, Springfield, Virginia 22151.		
19. Security Classif. (of this report) Unclassified	20. Security Classif. (of this page) Unclassified	21. No. of Pages 386	22. Price

Form DOT F 1700.7 (8-69)

16. **Abstract** *(Continued)*

(b) Conventional angle measurement techniques typically will give the desired performance only when interfering signals have an angular separation of at least 1 to 2 antenna beamwidths from the desired signal. Several improved signal processing techniques can reduce this minimum angular separation against coherent interference (e.g., multipath) by a factor of 3 to 5 if there is sufficient SNR (e.g., 30 dB) and an appropriate array configuration. Order of magnitude improvements in the minimum angular separation are achievable against incoherent interference (e.g., synchronous garble in ATCRBS and BCAS) with a moderate SNR (e.g., 20 dB). With either type of interference, time averaging of received wavefront data can significantly reduce the minimum SNR required over that for a small number of observations.

(c) The computations required for certain of the high performance algorithms seem readily implementable.

(d) Receiver calibration and monitoring will be a potentially significant factor in practical realization. Further study is needed to determine the required amplitude and phase accuracies.

(e) It is doubtful that these techniques can be easily applied to MLS angle processing due to the MLS signal format and the structure of the current MLS receivers.

ACKNOWLEDGMENTS

CONTENTS

ILLUSTRATIONS

ILLUSTRATIONS (cont.)

ILLUSTRATIONS (cont.)

ILLUSTRATIONS (cont.)

ILLUSTRATIONS (cont.)

ILLUSTRATIONS (cont.)

I. INTRODUCTION

A. Background

Historically, a principal method of improving the angle measurement accuracy of air traffic control (ATC) navigation aids and/or surveillance systems has been to increase the aperture of the ground antennas. The rapidly increasing capability and lower cost of digital signal processing hardware has made it possible to consider sophisticated signal processing techniques as an alternative method of improving system performance.

Some work in this area has already taken place. For example, all three of the techniques considered by the International Civil Aviation Organiation (ICAO) for adoption as the international standard Microwave Landing System (MLS) used novel signal processing algorithms to achieve better performance against multipath than would be possible with classical angle determination algorithms (see Table 1-1). More recently, it has been suggested that the DME based Azimuth System (DAS) could utilize the DME Based Landing System (DLS) array signal processing techniques to achieve the desired terminal area navigation performance at difficult sites with electrically small apertures.

Additional stimulus for considering performance improvement options arises from several current trends in navigation and surveillance system requirements:

(1) increased coverage to support noise abatement and/or fuel saving paths

(2) increased accuracy to support better airspace and airport utilization (e.g., closely spaced parallel runways and/or metering and spacing)

(3) increased usage at sites (e.g., oil well platforms, STOL sites, mountainous regions) where obstructions (i.e., multipath sources) may be much closer to the ground antenna than was the case in operations at major airport runways for conventional aircraft.

It should be noted that the system siting to achieve goals (1) and (2) frequently entails locating an antenna in a location (e.g., top of a building)

1-1

TABLE 1-1

NOVEL SIGNAL PROCESSING TECHNIQUES USED BY
MICROWAVE LANDING SYSTEM (MLS) PROPOSERS

TECHNIQUE	NOMINAL ANGLE OF ARRIVAL ESTIMATION TECHNIQUE	
DME based Landing System (DLS)	ground based multiple baseline interferometers	circular estimation algorithms adaptive subaperture beamsteering to minimize multipath spatial averaging or beamsteering in the orthogonal plane to reduce off-axis multipath levels
Doppler MLS (DMLS)	commutated source Doppler frequency estimation by sum/ difference correlator	randomized source commutation in the orthogonal plane to reduce off-axis multipath levels adaptive digital filtering to reduce flare multipath from the terrain
Time Reference Scanning Beam (TRSB)	centroid estimation of time between passages of a scanned beam	estimation of beam passage on the beam edge which is least perturbed by terrain reflection multipath

in which small physical size is very desirable. Achieving the requisite performance improvement by changing to a new form of navigation aid or surveillance system is increasingly difficult due to the economics of transition and the increasing demands on the frequency spectrum.

As noted in Table 1-1, there has been considerable work on advanced signal processing in the MLS context. Additionally, there has been extremely active recent work in related time series and array processing (e.g., sonar and seismic) areas [1, 2, 3] focused on high resolution spectrum estimation techniques. In the later approaches spectral features which normally would be obscured by the available time window [2] are recovered by improved signal processing techniques. The well-known duality between time series and line arrays, i.e.,

$$\text{time} \longleftrightarrow \text{spatial location within array}$$
$$\text{frequency} \longleftrightarrow \text{sine (angle of arrival)},$$

shows that much of the time series high resolution spectrum estimation work is applicable to the problem of estimating the direct signal angle of arrival in the presence of competing signals.

Many of the high resolution spectrum analysis studies have been ad hoc in nature and/or not considered key aspects of the ATC navigation/surveillance interference environment (especially multipath). Consequently, there has been an ongoing need for a more fundamental look at the possibilities offered by advanced signal processing that considers the principal navigation aid and surveillance system error sources from the outset.

B. Principal Error Sources

To achieve improved performance, one must reduce and/or better compensate for the principal error sources. The error sources considered in this study are:

(1) receiver front end noise

(2) signal waveform quantization (e.g., A/D errors)

(3) coherent interference (multipath), and

1-3

(4) calibration/monitoring errors (e.g., between various antenna elements in an interferometer array).

The four error sources above are not the only errors for a practical implementation (e.g., time quantization can be important for Time Reference Scanning Beam (TRSB) angle measurement, frequency stabilty for Doppler systems, computation round off errors for certain algorithms, etc.). However, these other factors typically reflect instrumental effects specific to a particular implementation as opposed to common major error sources. Some discussion is in order at this point regarding error sources (1) - (4).

1. Front End Noise Effects

Receiver front end noise[*] is a very important limitation both practically and theoretically. The system power budget is a key first step in any practical system implementation. Theoretically, this noise must ultimately limit the achievable accuracy with any processing scheme much as the noise power density limits the achievable capacity of a communication channel. An important aspect of the current study has been to make a start toward achieving an information theory type "channel capacity" for direction of arrival systems which would yield bounds on the theoretically achievable performance, thus yielding a "yardstick" by which one could assess the performance of practical signal processing techniques.

When the front end noise is the only error source, it can be shown [34, 42] that several common techniques (e.g., TRSB split gate trackers, the DAS linear interferometric arrays with a preponderance of elements at the array end points, and monopulse) closely approach the theoretical limits. Consequently, there is little to be gained in performance against front end noise by advanced signal processing techniques.

---

[*] and (for very low noise front ends) antenna temperature noise

2. Signal Waveform Quantization

Signal waveform quantization (e.g., A/D quantization errors) is often similar to front end noise in terms of its effects on the received waveform. It may, however, appear in different forms depending on the variables quantized (e.g., A/D conversion of amplitude and phase versus conversion of in-phase and quadrature components). Of course, A/D noise is an instrumental factor which could be reduced to an insignificant factor as technology progresses, however, the current state of technology and the dynamic ranges required in many applications is such that A/D quantization noise can be a practically important factor. Consequently, we have chosen to include it in the analysis.

3. Coherent Interference (Multipath)

Multipath has been found to be a major error source in most navigation and ATC surveillance systems – particularly when marginal or unsatisfactory performance occurs. Figure 1-1 shows the major multipath sources considered in the MLS assessment. In addition to those illustrated in Fig. 1-1, one must also consider shadowing by fixed objects (e.g., buildings or terrain).

Some perspective on the incidence of various phenomena at C band in the vicinity of an airport runway can be obtained from table 1-2 which summarizes the TRSB field tests at various operational runways. Similarly, the likely incidence of multipath phenomena at L band with 360° azimuth coverage can be inferred from the results of the Discrete Address Beacon System (DABS) Transportable Measurement Facility (TMF) tests at a variety of US sites. Severe reflection multipath conditions are typically manifested by false target reports. Figure 1-2 shows the incidence of false targets at the various TMF sites. The high incidence of false alarms at Washington National arose from building reflections whereas the false targets at Salt Lake City are due to reflections from a nearby mountain ridge [4].

The TMF tests also highlighted the problems in providing 360° azimuth coverage from a ground based site. Figure 1-3 shows the region where the TMF azimuth error exceeded the 0.1° objective as a function of azimuth and

1-5
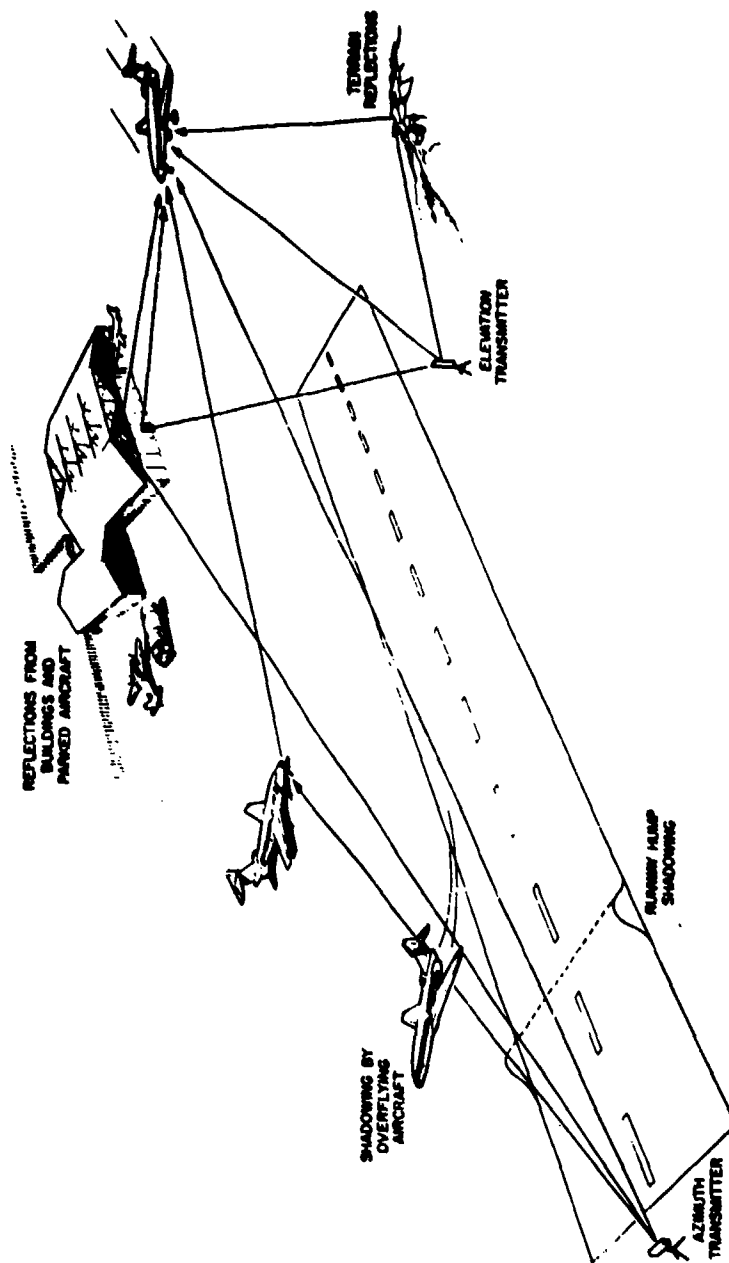
Fig.1-1. MLS multipath phenomena.

REFLECTIONS FROM BUILDINGS AND PARKED AIRCRAFT

TERRAIN REFLECTIONS

ELEVATION TRANSMITTER

SHADOWING BY OVERFLYING AIRCRAFT

RUNWAY HUMP SHADOWING

AZIMUTH TRANSMITTER

TABLE 1-2

MLS MULTIPATH DATA OBTAINED FROM OPERATIONAL AIRPORTS

Airport	Type of MLS System tested	Discernible Multipath Error Effects	
		Azimuth	Elevation
Buenos Aires Argentina	Basic Narrow	none	shadowing by ILS monitor
Hondouras	Small Community	small errors due to rough terrain reflections	
Kristiansand, Norway	Basic Narrow	small errors at low angles due to rough terrain shadowing and reflection	
Brussels	Basic Wide	very small	shadowing by C-130 aircraft
Charleroi, Belgium	Small Community	none or small errors due to terrain reflection/diffraction	
Dakar,Senegal	"	"	
Nairobi, Kenya	"	"	

TABLE 1-2 (continued)

MLS MULTIPATH DATA OBTAINED FROM OPERATIONAL AIRPORTS

Airport	Type of MLS System tested	Error Effects Azimuth	Elevation
Cape May, N.J.	"	"	
Crows Landing, California	Basic Narrow and Small Community	none	shadowing by monitor poles
J. F. Kennedy, N.Y.	Basic Narrow and Test Bed	none	reflections and/or shadowing by buildings
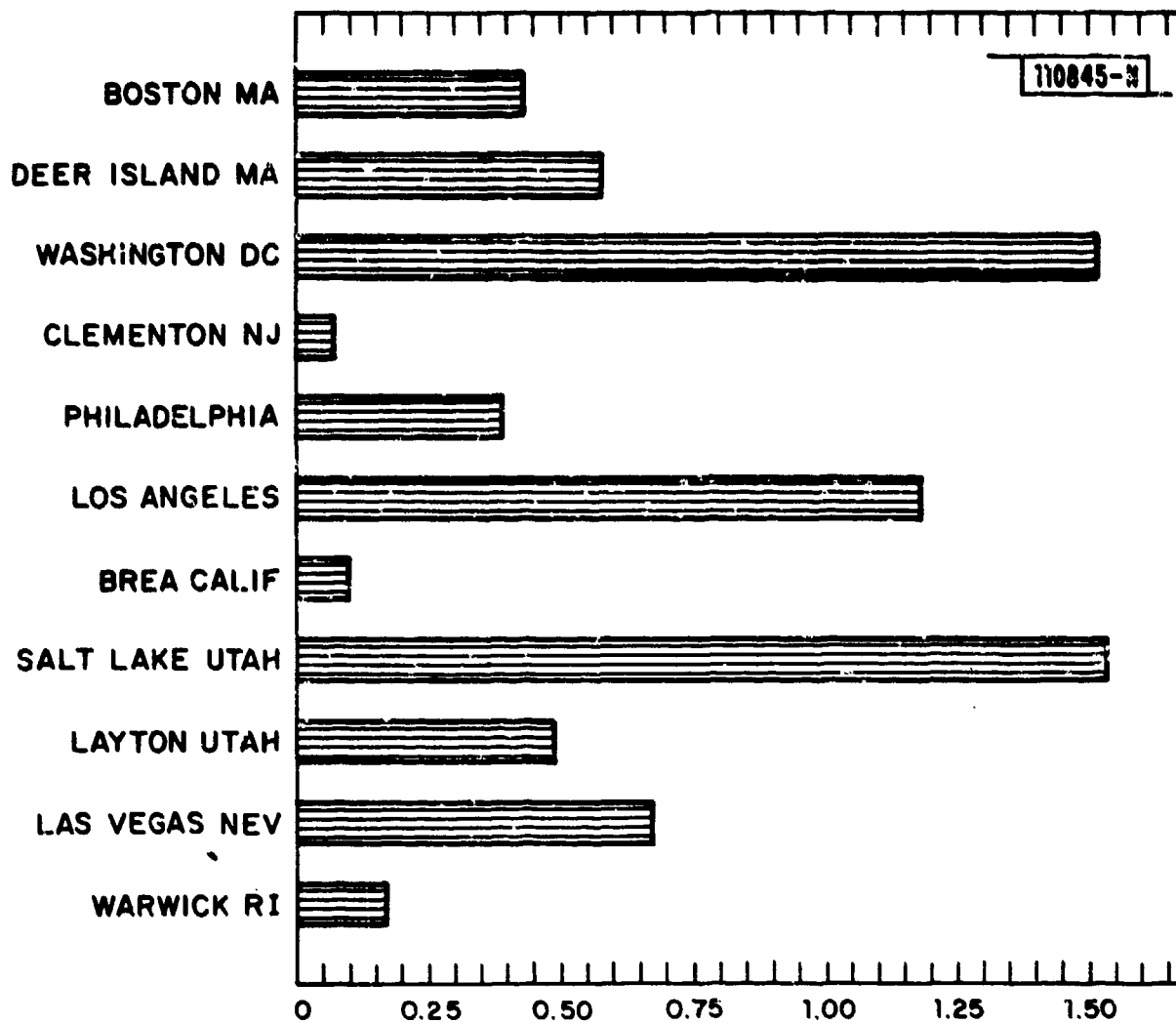
Fig.1-2. Reports labeled as false (percent of all reports).
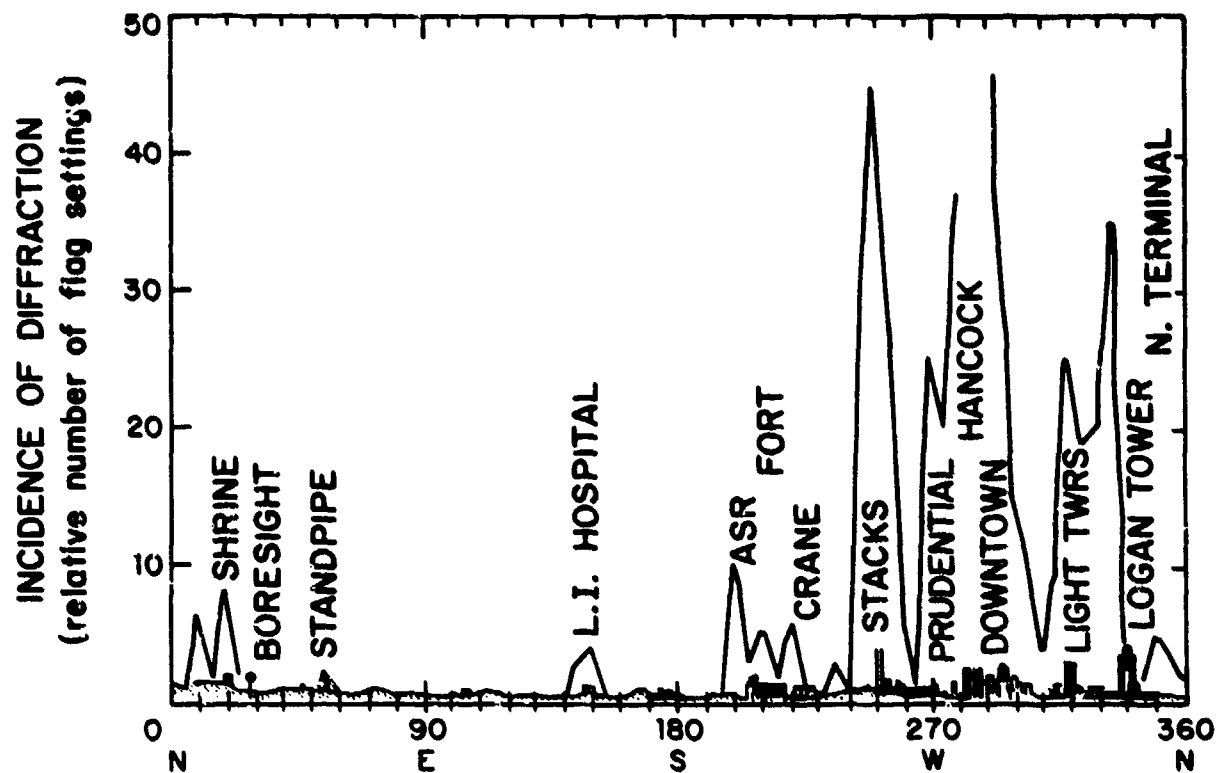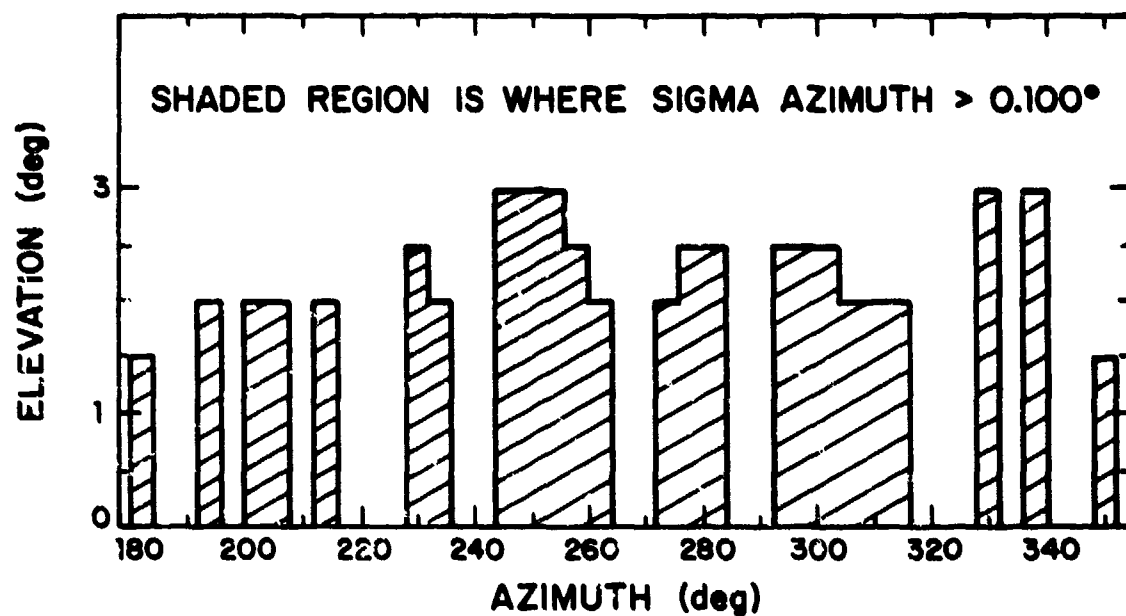
Fig. 1-3. Incidence of monopulse diffraction errors at Logan Airport.

elevation angles together with the local skyline at Logan International Airport (Boston). We see that the regions of unacceptable performance correspond closely to diffraction by buildings, and other skyline objects. Similar diffraction phenomena were encountered at many of the other TMF sites [6-8].

The incidence of diffraction phenomena at low elevation angles is of particular importance for navigation and surveillance systems which provide enroute or terminal area $360°$ azimuth coverage. Studies of the distribution of aircraft as a function of elevation angle show that most of the aircraft within 150 miles of a ground site are at elevation angles less than $2°$ [70]. Figure 1-4 shows one such computed distribution. We see that 25% of the aircraft are at an elevation angle less than $1°$.

One of the current ATC navigation deficiencies is the provision of high quality navigation data in mountainous regions. In fact, the need for such information has been an important stimulus for developing a $360°$ azimuth capability in the MLS system [41]. Sites with upsloping and irregular terrain present a multipath challenge to azimuth, elevation and ranging systems since the multipath location parameters (e.g., relative azimuth, elevation, time delay) can have a far greater spread than is the case over flat terrain. It had been thought that the small scale roughness associated with such terrain would cause the multipath levels to be significantly lower than was the case for nominally flat terrain. However, recent measurements over irregular surfaces at Camp Edwards, Mass., (Figs. 1-5 and 1-6) have shown (Fig. 1-7) that such terrain can yield specular multipath with levels comparable to the reflections which arise over flat terrain [5].

Recent work in multipath modeling [5, 6] has shown that both reflection and diffraction phenomena can be represented by a sum of discrete rays (i.e., plane waves) which have essentially the same characteristics* as the direct signal ray from transmitter to receiver. Improved signal processing

---

*e.g., the multipath signals have the same signal waveform as the direct path signal when the receiver is not moving.
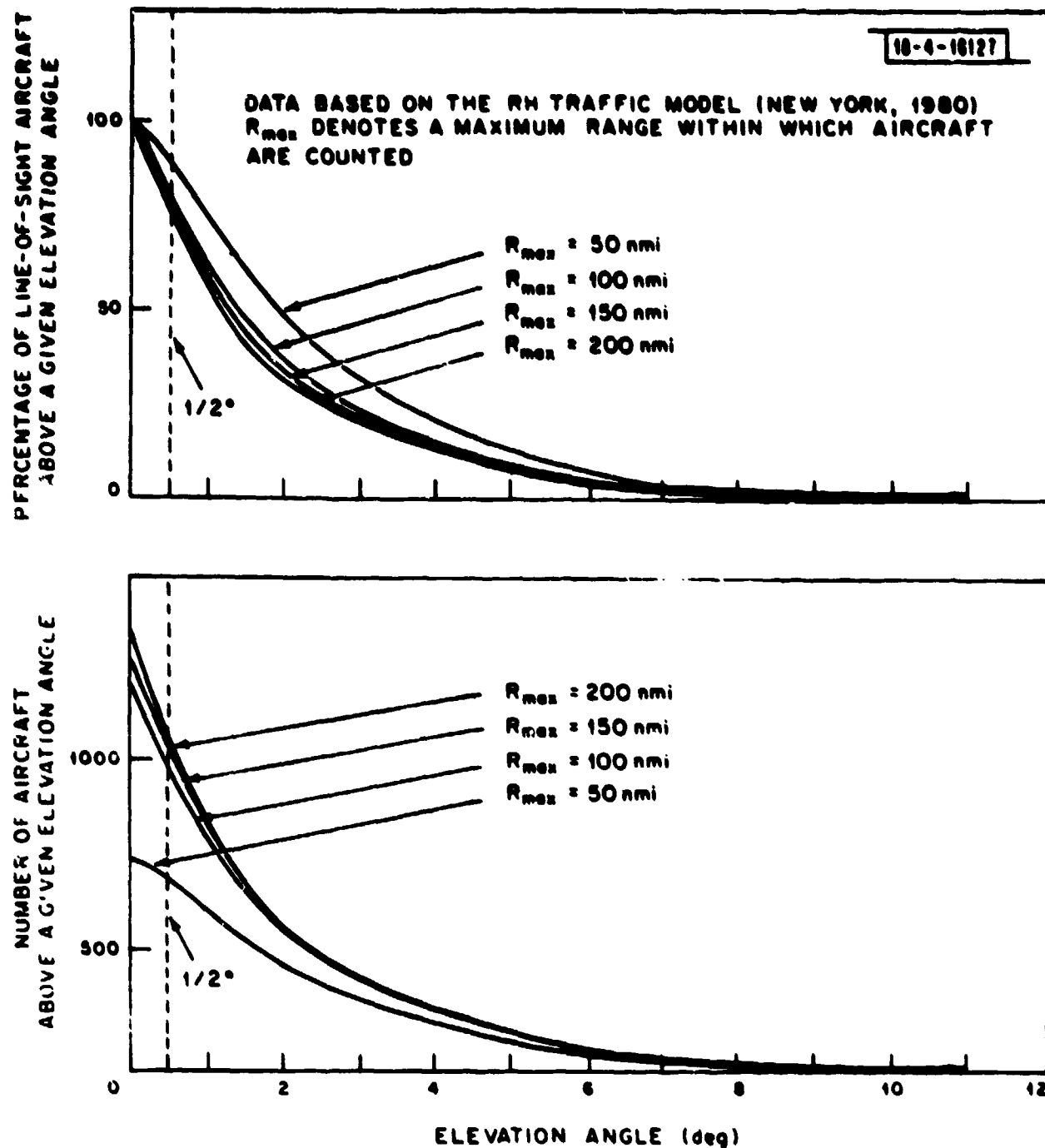
Fig. 1-4. Cumulative distribution of aircraft count with elevation angle.

Fig. 1-5. Terrain along road at Camp Edwards Mass.

DISTANCE FROM ANTENNA (FEET)
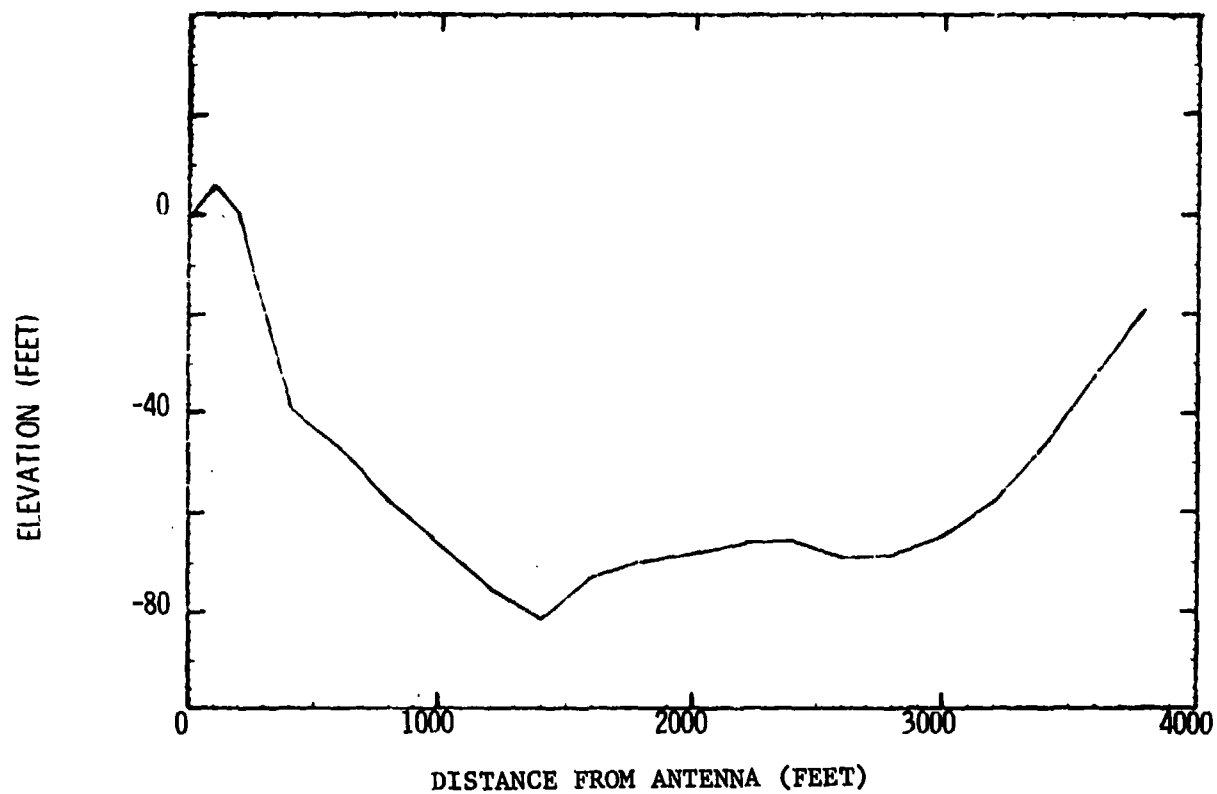
Fig. 1-6. Terrain profile at Camp Edwards, Mass. Site #2
(Gibbs Road)

1-13
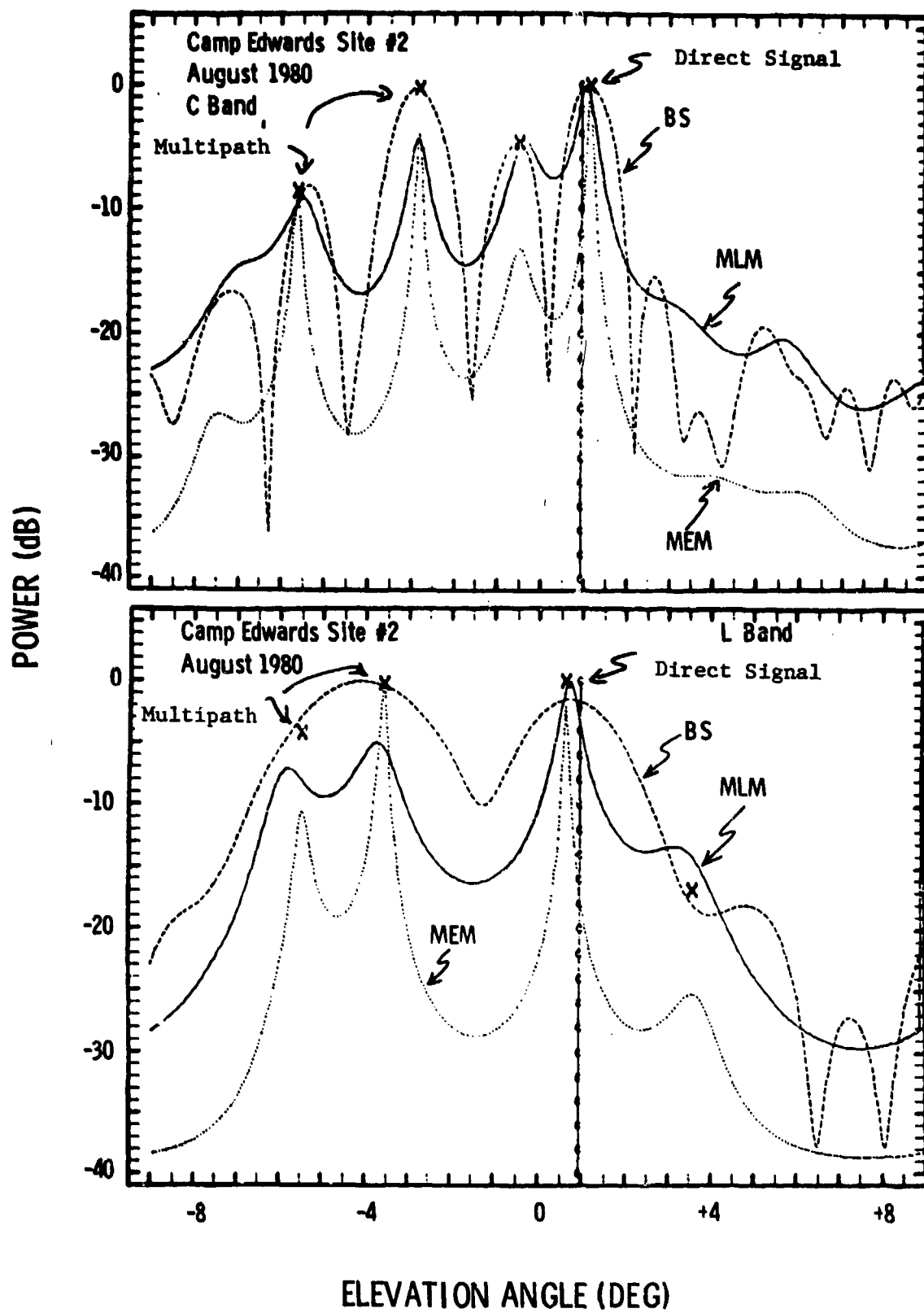
Fig. 1-7. Received Power Vs. Elevation Angle At
Camp Edwards Site #2

1-14

techniques must capitalize on the (often small) differences between multipath rays and the direct signal ray in terms of

(1) azimuth and elevation angle as measured from the ground antenna

(2) time difference

(3) amplitude

(4) carrier frequency, and

(5) relative rf phase

For angle (i.e., azimuth or elevation) measurement navigation or surveillance systems, a key issue is whether the multipath signals are inbeam or out of beam, i.e., whether the angular separation of the multipath signal from the direct signal is greater than 1.5 λ/D where D/λ is the ground antenna aperture in the scan plane[*]. Multipath which is out of beam typically can be handled with by more or less standard signal processing methods (i.e., low sidelobes and appropriate acquisition/validation algorithms [28, 29]). By contrast, inbeam multipath generally presents a much more difficult problem for standard signal processing approaches.

Angle errors due to diffraction generally are caused by 1 or 2 inbeam multipath rays [28] as are certain terrain reflection error phenomena. By contrast, building reflections are generally out of beam for azimuth measuring systems, but may be inbeam for elevation [29] angle measuring systems. As a consequence of the above considerations, methods of reducing the effects of inbeam multipath has been a principal consideration in the studies reported here.

---

[*]for ranging (e.g., DME) systems, the corresponding condition is multipath with a time delay $\leq$ 1/signal bandwidth.

4. Calibration/Monitoring Effects

Many of the advanced signal processing methods of greatest interest typically attempt to estimate additional features (e.g., error sources) of the received signal environment and then utilize that information in estimating the direct signal parameters. Others attempt to use certain details in the received signal structure to favor the desired signal*. In either case, system performance may become much more sensitive to certain system imperfections such as signal path equalization and component operation. Consequently, the degree to which one must calibrate and monitor the various signal paths is an important practical consideration in considering advanced signal processing techniques.

5. Summary of Error Source Discussion

From the considerations above we concluded that improving multipath performance should be a principal objective for advanced signal processing techniques since the other error sources are typically 1) not the primary limitation on achievable performance or 2) (as in the case of front end noise) there is theoretical reason to believe that little improvement is possible over current processing approaches.

C. Objectives of This Study

The overall objective of the study reported here has been the evaluation of signal processing techniques for navigation applications utilizing data processing as a means of improving system accuracy without increasing the antenna size. Since most of the results are also applicable to ATC surveillance systems, we have not tried to separately address the case of surveillance systems. The principal focus in the work has been on methods which accomplish wavefront analysis with a multiple element/multiple receiver array (such as shown in Fig. 1-8) since:

---

*e.g., MLS flare guidance "single edge" processors, DME delay and compare circuits [28].

1-16

PHASE SHIFTERS

z
z^2
z^3
z^4
z^5

OUTPUT WHEN POINTED
AT ANGLE ϕ

$z = \exp(jA \sin \phi)$

ϕ
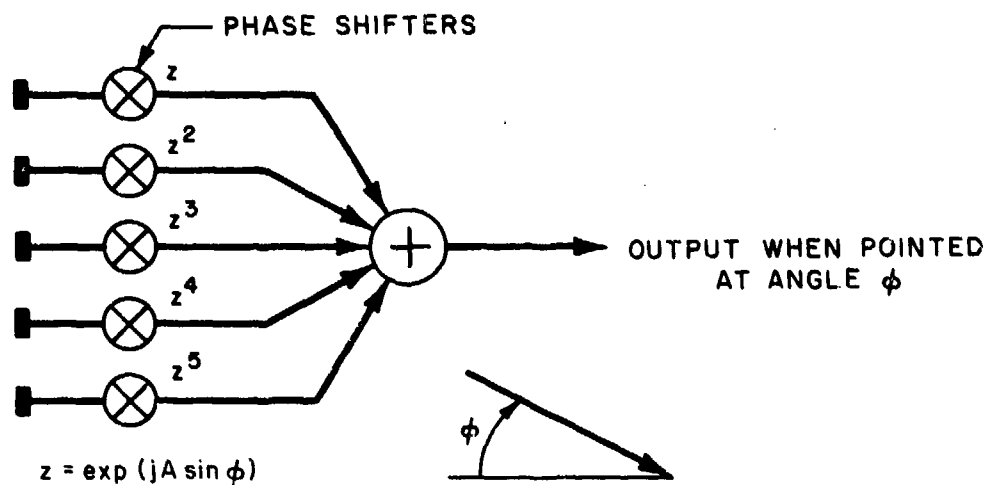
CONVENTIONAL ANTENNA BEAMFORMING

A/D
A/D
A/D
A/D
A/D

COMPUTER

ESTIMATE OF POWER
COMING FROM ANGLE ϕ
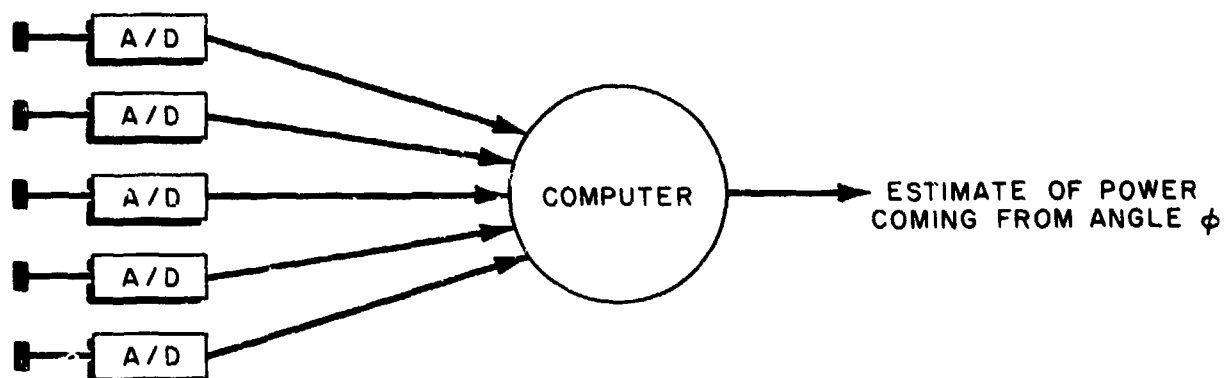
SAMPLED APERTURE PROCESSING

Fig. 1-8. Comparison of sampled aperture processing with conventional beamforming.

(1) this type of system preserves the wavefront information (e.g., phase and amplitude) along the aperture

(2) the bulk of recent algorithms and theory utilize such a system as their starting point

(3) most air derived navigation systems can be shown (by reciprocity considerations) to be equivalent to such a system, and

(4) the DAS system under study as a means of providing 360° terminal area azimuth coverage in an FAA/Federal Republic of Germany (FRG) joint program utilizes such an array.

Each of the points [(1) - (4)] above will be detailed at various points in the subsequent chapters.

The array signal processing studies have emphasized methods for improving multipath performance. Specifically, we have sought to:

(1) analytically calculate the (theoretically) achievable performance with particular emphasis on the effects of front end noise and/or equipment errors (e.g., calibration errors, and A/D quantization) in the presence of multipath,

(2) assess the comparative performance of candidate advanced array signal processing techniques including FRG scheme(s) and representative US technology (e.g., maximum likelihood methods and autoregressive modelling) via analytic studies, simulations and sampled aperture field data,

(3) understand hardware implications of practical implementation for the more promising techniques (e.g., cost and complexity to achieve desired performance and capacity), and

(4) investigate how the aperture sampling technique results might be applied to TRSB angle guidance signal processing.

D. Report Outline

The model for the received signal characteristics is a key step in the development of improved signal processing techniques. Chapter II discusses appropriate models for a single received wavefront and a time sequence of

1-18

wavefronts. These models include both deterministic signal waveforms (with unknown parameters) and random process representations whereby the observed data is the output of a deterministic linear filter (with unknown parameters) being driven by a random process.* Also discussed in chapter II is the concept of a spatial covariance function since the use of such covariances are a key feature of many advanced signal processing techniques.

Chapter III discusses theoretical performance bounds on estimating the direct signal parameters (especially, angle in the scan plane of the array) when front end noise is present. These bounds provide benchmarks to compare the performance of the various specific techniques which are discussed in chapter IV. The Chapter IV technique discussion includes methods suggested by the FRG for use with DLS/DAS as well as a number of techniques under active investigation in the U.S.

Section V quantifies the performance of a number of the most promising techniques at mitigating the errors due to multipath (principally inbeam). The comparison here utilizes both simulation results and the results of field measurements with several experimental arrays. It is shown that when the received wavefront complexity is low (i.e., a small number of received rays) and reasonable signal to noise ratios (SNR) exist, certain advanced signal processing techniques can provide a substantial increase in performance over classical approaches. With complicated wavefronts and/or low SNR, the advanced method performance typically is comparable to that of classical methods.

In view of the direct applicability of the advanced techniques to DAS type aperture sampling systems, an investigation of unique hardware implementation issues for such a system was conducted. Section VI discusses a number of these issues in the context of the Lincoln experience in using such a system to analyze terrain reflection multipath and experimentally validate the simulation results for multipath rejection by advanced signal processing techniques.

- - - - - - - - - - - - - -

*This latter model is similar to that which is utilized in Kalman filtering

Section VII considers the application of the advanced signal processing techniques to the MLS angle guidance or subsystems. Although reciprocity/duality considerations would allow one to utilize such techniques, at least in principal, there are practical problems associated with the loss of phase information in envelope formation. Two avenues around this problem are considered: coherent demodulation (using, e.g., the TRSß DPSK transmission) and direct reconstruction of aperture information from the envelope data.

The final section summarizes the results and makes several recommendations for future work.

II. ALGEBRAIC DIRECTION FINDING PROPERTIES OF LINEAR ARRAY ANTENNAS

This chapter is concerned with certain fundamental limitations on the direction finding capability of a linear array. Our basic assumption is that the signal environment is specular; i.e., the array essentially sees only a finite number of plane waves. Specifically, we consider ideal arrays of uniformly spaced, isotropic elements. In principle, our basic model can be extended to include more realistic situations. For example, nonuniform arrays and/or mutual coupling between elements [71] can be handled by specifying the correct functional form for the direction (steering) vectors. However, the reader is advised to exercise a fair amount of prudence in applying our results to non-ideal situations.

As will be seen in Chapter III, the resolution capability of an array is theoretically limited by the total available signal-to-noise ratio. In practice, most "super-resolution" algorithms (see Chapter IV) require many time observations of the array signals in order to perform well.

The treatment of "multiple look" algorithms is simplified by adopting a suitable stochastic model for the received signals. We use the standard complex representation for the observed quadrature data, and the signal and noise components of the data are assumed to be uncorrelated circular (complex) random processes with zero mean. Unless otherwise stated, the noise is "white"; i.e., the noise covariance matrix $\underline{N}$ is a (positive) scalar multiple of the identity matrix $\underline{I}$.

Since many fundamental results depend only upon second-order statistics, a significant fraction of this chapter is devoted to a study of the algebraic properties of the signal covariance matrix $\underline{S}$. An attempt has also been made to motivate certain covariance estimates that have generally proven useful in direction-finding applications. These estimates are described in terms of formal matrix operations applied to sample covariance matrices.

A. Signal and Covariance Models

In subsection 1, we present the basic signal model and briefly review an exact method, due to Prony [48], for extracting directional information from

ideal data. Unfortunately, when a modest amount of noise is present, the direction-finding performance of the (exact) Prony method and similar approaches may be quite poor. After introducing appropriate statistical models for signal and multiple time observations in subsection 2, we consider a least squares extension of Prony's method in subsection 3. This approach leads naturally to the concept of a time-averaged and spatially smoothed covariance estimate. Time-averaged covariance estimates (i.e., sample covariance matrices) are frequently encountered in the contex of (digital) adaptive arrays, but the role of spatial smoothing in direction-finding was not (widely) recognized until relatively recently. The relationship between the adaptive array theory and the covariance estimation techniques which arise with the extended Prony and certain other methods is emphasized in subsection 4.

1. Ideal System Model

The postulated signal environment on a single observation, consists of I narrowband plane waves arriving from distinct directions θ_i. The complex amplitude of the ith signal at the array phase center is p_i, and the signal received at the ℓth element is

$$s_\ell = \sum_{i=1}^{I} p_i \exp \{j\alpha_i[\ell-(L+1)/2]\}; \quad \ell=1, \ldots L \quad . \quad (2.01)$$

The received signals depend on the directions of arrival only through the (phase) angles

$$\alpha_i = 2\pi \frac{\xi}{\lambda} \sin \theta_i; \quad i=1, \ldots I \quad (2.02)$$

where ξ is the element separation and λ is the (common) wavelength of the signals in space. For arbitrary α, we introduce the L-dimensional direction vector $\underline{v}_L(\alpha)$ with elements

$$v_\ell(\alpha) = \exp \{j\alpha[\ell-(L+1)/2]\}; \quad \ell=1, \ldots .L \quad . \quad (2.03)$$

Except when needed to avoid confusion, the (length) subscript L for a direction vector is usually omitted.

The basic linear relationship (2.01) between the signals received by the elements and the complex signal amplitudes at the array phase center can be conveniently expressed in standard matrix notation,

$$\underline{s} = \underline{V}\, \underline{p} \qquad , \tag{2.04}$$

where $\underline{V} = \underline{V}(\underline{\alpha})$ is easily constructed from the true direction vectors, i.e.,

$$\underline{V}(\underline{\alpha}) \overset{\Delta}{=} [\underline{v}(\alpha_1) \cdots \underline{v}(\alpha_I)] \tag{2.05}$$

The L-dimensional vector $\underline{s}$ is called the received signal. To avoid possible confusion, we will refer to the I-dimensional vector $\underline{p}$ as the signal-in-space.

An attractive way of estimating signal directions is to examine the angular spectrum of the received signal. Conventional spectral (periodogram) analysis suggests that we calculate the complex inner product

$$S(\alpha) \overset{\Delta}{=} \underline{v}^H(\underline{\alpha})\, \underline{s} \tag{2.06}$$

where the subscript H denotes the complex conjugate (Hermitian) transpose. Angle estimates are then extracted from the spectral peaks, i.e.,

$$\underset{\alpha}{\text{maxima}}\ |S(\alpha)| \implies \{\hat{\alpha}_i\} \qquad .$$

When only one signal-in-space exists, the angle spectrum is easily computed. Setting $\underline{s} = \underline{v}(\alpha_1)$ in (2.06) gives

$$S_1(\alpha) \overset{\Delta}{=} \underline{v}^H(\alpha)\, \underline{v}(\alpha_1)$$
$$= w(\alpha - \alpha_1)$$

where

$$w(\alpha) \overset{\Delta}{=} \frac{\sin \alpha L/2}{\sin \alpha/2}$$

is the familiar "gain" pattern characteristic of a uniformly illuminated aperture. The first null of $w(\alpha)$ occurs at

$$\alpha_B \overset{\Delta}{=} 2\pi/L \qquad\qquad (2.07)$$

which suggests that α_B is the appropriate definition of "angular" beamwidth.

The incremental relationship

$$\Delta\alpha = 2\pi \frac{\xi}{\lambda} \cos\vartheta\ \Delta\theta$$

between the _angle_ α and the _direction_ θ follows easily from (2.02). Rearranging this equation leads to

$$\frac{\Delta\alpha}{\alpha_B} = \Delta\theta\ \frac{L\xi\ \cos\theta}{\lambda}$$

which establishes the equivalence between α_B and the directional beamwidth,

$$\theta_B \overset{\Delta}{=} \frac{\lambda}{L\xi\ \cos\theta} \qquad\qquad (2.08)$$

conventionally used for arrays with aperture length $L\xi$. The $\cos\theta$ factor in (2.08) is explained by the effective loss of aperture exceeded by a (broadside) array when receiving a signal in any direction other than $\theta = 0$.

When more than one signal in space exists, the angle spectrum based on a single periodogram is specified by a sum of $(\sin Lx)/\sin x$ functions

$$S(\alpha) = \sum_i P_i\ w(\alpha - \alpha_i)$$

When all of the signals present are separated by two or more beamwidths, the magnitude of $S(\alpha)$ will have separate peaks which can be associated with the various signals in the sense that the peak locations are close to the actual angles of arrival $\{\alpha_i\}$. For angular separations between one and two beamwidths, separate peaks may or may not occur depending on the relative amplitudes and phases of the signals. When the dominant peak locations

coincide with the actual angle of arrivals, we say that the signals are resolved. Normally, two signals separated by less than a beamwidth cannot be resolved with Fourier (i.e., periodogram) methods. In fact, there is no single minimum separation for resolution applicable to all values of the signal-in-space (complex) amplitudes $\{p_i\}$. However, the conventional practice is to use the Rayleigh criterion* of one beamwidth separation as an indication of resolution capability for conventional array processing. The periodogram spectrum estimate is essentially identical to TRSB envelope observation. In the TRSB context, lack of resolution is typically associated with "inbeam" multipath. The TRSB rule of thumb has been to use an angular separation of 1.5 beamwidths as the inbeam/out-of-beam boundary.

The desire for improved resolution has led to a number of adaptive "super-resolution" algorithms based on the same ideas found in <u>nonlinear</u> spectral estimation methods (e.g., see [19]). An algorithm theoretically capable of "super-resolution" first appeared nearly two centuries ago. Generally known as Prony's method [45], the algorithm has recently been discussed in [46]. A description of Prony's algorithm is also available in [47].

Prony's Method

To apply the exact Prony algorithm, the signal data are assumed to satisfy a linear, recursive relationship of the form

$$s_\ell + \sum_{i=1}^{I} a_i \, s_{\ell+i} = 0 \qquad ; \qquad \ell = 1, \ldots, L - I \qquad . \qquad (2.09)$$

The complex <u>resonances</u> of the signal are the roots $\{z_i\}$ of the polynomial

$$A(z) \overset{\Delta}{=} 1 + a_1 z + \ldots + a_I z^I \qquad\qquad (2.10)$$
$$= (1 - z/z_1) \ldots (1 - z/z_I) \qquad .$$

_____

*The Rayleigh criterion is based on the spectrum obtained from averaging a number of observations, as will be discussed in the next section.

One first solves an appropriate set of I equations obtained from (2.09) for the unknown polynominal coefficients $\{a_1, \ldots a_I\}$ and then computes the roots of the polynomial (2.10) to determine the signal resonances.

To show that any signal of the form

$$s_\ell = \sum_{i=1}^{I} c_i z_i^\ell \qquad ; \qquad \text{arbitrary } \ell \tag{2.11}$$

can be represented by (2.09), first construct the polynomial coefficents from (2.10) using the resonant "frequencies" $\{z_i\}$ in (2.11) as the roots of the polynomial. Consequently, $A(z_i)$ is zero by construction, and hence we can immediately write

$$\sum_{i=1}^{I} c_i z_i^\ell \left(1 + \sum_{m=1}^{I} a_m z_i^m\right) = \sum_{i=1}^{I} c_i z_i^\ell A(z_i) = 0$$

A simple rearrangement of the left-hand side leads to

$$\left[\sum_{i=1}^{I} c_i z_i^\ell\right] + \sum_{m=1}^{I} a_m \left[\sum_{i=1}^{I} c_i z_i^{m+\ell}\right] = 0$$

Simplifying the bracketed terms with (2.11) yields the desired result (2.09).

Comparing (2.11) with (2.01), the resonant frequencies in the direction finding application considered here are seen to lie on the unit circle, i.e.,

$$z_i = e^{j\alpha_i} \qquad . \tag{2.12}$$

Consequently, the angles of arrival can be easily obtained via

$$\alpha_i = \arg z_i \qquad . \tag{2.13}$$

Once the correct frequencies are known, one solves an appropriate set of equations from (2.11) for the remaining coefficients $(c_1, \ldots c_I)$; together with the angles $\{\alpha_i\}$, these coefficients determine the complex amplitudes $\{p_i\}$ of the signals-in-space. When the actual data samples do not satisfy (2.09)

due to additive noise, etc., the estimates generated by the exact Prony algorithm may be significantly in error.

2. Received Signal Models

In practice, the observed signal is nearly always modeled by superimposing a "noise" vector on the signal actually received. The discussion in this chapter is based on the premise that the dominant sources of system error exhibit the same temporal characteristics as additive thermal noise.

Assuming K independent observations are made, the data available for processing are represented by

$$\underline{r}(k) = \underline{s}(k) + \underline{n}(k) \quad ; \quad k=1, \ldots K \quad . \tag{2.14}$$

Each vector of observed data is referred to as a "snapshot," and the elements of a snapshot are called its "samples." Thus, the available data generally consist of K snapshots, taken at more or less arbitrary instants of time. Each snapshot contains L samples, taken simultaneously at the uniformly spaced elements of the array. The noise component of the data, $\underline{n}(k)$, is assumed to be a white, ergodic random process with respect to the snapshot (time) index k.

Over typical processing intervals, the directions of arrival will not change significantly. In contrast, the signals-in-space typically vary with time. Since the behavior of the signal source(s) is generally unpredictable, a stochastic signal model is often appropriate. In this case, the statistics of the signal process are usually assumed to be time invariant. In general, we write the covariance (matrix) of an arbitrary complex random vector $\underline{z}$ as

$$\underline{z} \overset{\Delta}{=} E \{\underline{z} \, \underline{z}^H\} \tag{2.15}$$

where E{ } is the expected value or "ensemble average" operator for the underlying probability space.

When the signal and noise processes are essentially uncorrelated, the covariance relationship

$$\underline{R} = \underline{S} + \underline{N} \qquad (2.16)$$

follows directly from (2.14). The relationship between the (received) signal covariance $\underline{S}$ and the signal-in-space covariance $\underline{P} \triangleq E\{\underline{p}\,\underline{p}^H\}$ follows easily from (2.04), i.e.,

$$\underline{S} = \underline{V}\,\underline{P}\,\underline{V}^H \quad . \qquad (2.17)$$

The diagonal elements of $\underline{P}$ represent the power levels of the signals-in-space. If the off-diagonal elements are identically zero, then the signals-in-space are said to be (jointly) uncorrelated and, for an ideal array, the signal covariance matrix $\underline{S}$ will exhibit the Toeplitz property, i.e., $S_{mn} = C_{m-n}$. Any signal $\underline{s}$ with a Toeplitz covariance matrix is said to be spatially stationary. We generally refer to a Toeplitz covariance matrix as a correlation matrix.

The stationary case occurs most naturally when the signals-in-space originate from (statistically) independent sources. Examples of uncorrelated sources for ATC application include:

(1) overlapping DME interrogations at a DAS ground station,

(2) synchronous and asynchronous garble at an ATCRBS or BCAS angle-of-arrival sensor,

(3) specular and/or diffuse multipath which has a large relative Doppler frequency (i.e., scalloping rate) compared to the system measurement rate.

If significant specular reflection or diffraction multipath with a low scalloping rate is present, the signals-in-space are no longer uncorrelated since their relative phases and amplitudes are determined by geometry and the physical environment (terrain features, buildings, etc.). When multipath is inbeam, the signals-in-space are likely to maintain a fixed amplitude and phase relationship over an entire processing interval. In this case, the

signals-in-space are said to be <u>perfectly</u> correlated and $\underline{P}$ is a singular matrix.

3. Estimation of Covariance Functions

When system errors are present, one can proceed by several different avenues to estimate the signal parameters.

(1) estimate the plane-wave parameters on individual snapshots using (2.01) and (2.14). This approach leads one to methods such as the maximum likelihood estimation procedure to be discussed in the next chapter.

(2) utilize a procedure which is based on knowledge of the ensemble signal and noise covariances. The ensemble covariance must be estimated from the sample data. A variety of techniques exist based upon various covariance estimates, and several of these will be discussed in Chapter IV.

(3) make a least-squares estimate of the parameters in a signal model which encompasses the range of expected signal environments. This typically leads to an explicit form of covariance estimation.

Here, we consider a least-squares extension to the exact Prony method. This approach leads to a robust covariance estimation technique that has been proven effective in a wide variety of situations and also leads naturally to the concept of spatial averaging.

When random noise is present, the basic recursive relationship (2.09) is only approximately true for the <u>observed</u> samples. A pragmatic approach that leads to a very tractable algorithm is to choose polynomial coefficients that minimize the deviations from the ideal Prony model. For example, by introducing the error at the ℓth element as a function of the time index k. i.e.,

$$e_\ell(k) \overset{\Delta}{=} r_\ell(k) + \sum_{i=1}^{I} a_i r_{\ell+i}(k) \quad , \quad (2.18)$$

we can easily find the coefficients that yield the smallest errors over $M = L-I$ elements and K snapshots in the usual least-squares sense.

To solve this problem effectively, introduce

$$\underline{w}^H \triangleq (1, a_1, \ldots a_I) \tag{2.19}$$

and

$$\underline{r}_m(k) \triangleq [r_m(k), r_{m+1}(k), \ldots r_{m+I}(k)]^T \tag{2.20}$$

where a superscript T denotes the usual transpose operation. Using these definitions, we may write (2.18) as

$$e_\ell(k) = \underline{w}^H \underline{r}_\ell(k) \quad .$$

Thus, the minimum mean square error

$$\sigma^2 \triangleq \min_{\{a_i\}} \frac{1}{MK} \sum_{m=1}^{M} \sum_{k=1}^{K} |e_m(k)|^2 \tag{2.21}$$

is completely equivalent to

$$\sigma^2 = \min_{\underline{w}} \underline{w}^H \left[\frac{1}{MK} \sum_{m=1}^{M} \sum_{k=1}^{K} \underline{r}_m(k) \, \underline{r}_m^H(k) \right] \underline{w} \tag{2.22a}$$

subject to the constraint

$$w_1 = 1 \quad . \tag{2.22b}$$

The problem in (2.22) is that of minimizing a quadratic form with a (simple) linear constraint on the weights. The general solution appears later in a slightly different context. For the present, our interest lies in the operations to be performed on the data.

The data-dependent term appearing in (2.22a) can be expressed in terms of the usual "sample" covariance matrix

$$\underline{\hat{R}}(K) \triangleq \frac{1}{K} \sum_{k=1}^{K} \underline{r}(k) \, \underline{r}^H(k) \tag{2.23}$$

by introducing the rectangular sampling matrices of the form

$$\underline{A}_m = (\underline{0} \mid \underline{I} \mid \underline{0}')$$ (2.24)

where the subscript m denotes the first column in which a "one" appears. The sampling matrices are useful for expressing the data vectors introduced in (2.20) in terms of the actual snapshot data, i.e.,

$$\underline{r}_m(k) = \underline{A}_m \, \underline{r}(k) \qquad .$$

Since $\underline{r}_m(k)$ is separable in m and k, it follows easily that

$$\frac{1}{MK} \sum_{m=1}^{M} \sum_{k=1}^{K} \underline{r}_m(k) \, \underline{r}_m^{H}(k) = \frac{1}{M} \sum_{m=1}^{M} \underline{A}_m \, \hat{\underline{R}}(K) \, \underline{A}_m^{T} \qquad .$$

The "spatial smoothing" operation that remains can be formally extended to any **square** matrix. As used previously, the sampling transformation $\underline{A}_m \underline{x}$ extracts a subcolumn of the elements of (a vector) $\underline{x}$. It follows that $\underline{A}_m \, \underline{H} \, \underline{A}_m^{T}$ is a **principal** submatrix of any (square) matrix $\underline{H}$. Therefore,

$$\underline{H}_M \overset{\Delta}{=} \frac{1}{M} \sum_{m=1}^{M} \underline{A}_m \, \underline{H} \, \underline{A}_m^{T}$$ (2.25)

is an "internal" average of the principal submatrices of $\underline{H}$; the subscript specifies the number of submatrices to average, and the submatrices must all have the same **maximal** dimension consistent with M. Of course, $\underline{H}_1 = \underline{H}$ and the subscript may be omitted.

With the help of the spatial smoothing operation, we may now express the original least-squares problem as

$$\sigma^2 = \min_{\underline{w}} \underline{w}^{H} \, \hat{\underline{R}}_M(K) \, \underline{w}$$ (2.26a)

subject to

$$w_1 = 1 \quad .$$
(2.26b)

Clearly, the fundamental operations of time-averaging and spatial-smoothing are interchangeable. These operations always yield at least a positive semidefinite Hermitian matrix when applied to arbitrary data.

The approach just described for estimating (Prony's) polynomial coefficients from noisy data is known as the covariance method [20]. This approach is intimately related to the ideas of linear prediction and autoregressive analysis. In fact, the covariance method may be viewed as an extension of Prony's method in which "autoregressive" coefficients are chosen to minimize the "prediction" errors by the traditional method of least squares.

4. Relation to Theory of Adaptive Arrays

Adaptive array theory is usually presented in terms of a statistical interpretation of the least-squares concept. In typical "nulling" problems, one attempts to minimize the expected "interference" power in a linear combination of the array samples, subject to a "steering" constraint on the element weights [49]. Thus, a fairly common adaptive array problem is to find a set of weights $\underline{w}$ that achieves the minimum expected output power

$$\sigma^2 \stackrel{\Delta}{=} \min_{\underline{w}} E\{|\underline{w}^H \underline{r}|^2\}$$
$$= \min_{\underline{w}} \underline{w}^H \underline{R} \, \underline{w}$$
(2.27a)

subject to a linear constraint, say

$$\underline{w}^H \underline{c} = 1$$
(2.27b)

where $\underline{R}$ is the covariance of the observed vector $\underline{r}$. A condition such as (2.27b), or some alternative constraint, is necessary to prevent $\underline{w} = \underline{0}$.

Several of the super-resolution techniques to be discussed in Chapter IV differ principally by the type of steering constraint applied to the weights. In general, the solution to (2.27) is

$$\sigma^2 = (\underline{c}^H \underline{R}^{-1} \underline{c})^{-1} \quad , \tag{2.28}$$

and the optimum weights are given by

$$\underline{w} = \sigma^2 \underline{R}^{-1} \underline{c} \quad . \tag{2.29}$$

In adaptive arrays configured for digital operation, the "true" covariance matrix $\underline{R}$ is often estimated from a sample covariance matrix, $\hat{\underline{R}}$ (K), as defined in (2.23). The ergodic assumption,

$$\lim_{K \to \infty} \hat{\underline{R}} (K) = \underline{R} \quad , \tag{2.30}$$

is (nearly) always valid, and hence essentially optimum nulling performance can be anticipated as the number of snapshots becomes sufficiently large.

Most adaptive array applications are primarily concerned with suppressing interference in order to enhance a "friendly" signal, usually assumed to be absent when the adaptive weights for the array elements are calculated. Until recently [50], the complementary problem of estimating the directions of the interfering sources had received relatively little attention in the adaptive array literature. The difficulties that arise with nulling in the coherent case were discussed by Gabriel [1], but the performance limitations imposed by multipath were not addressed quantitatively.

B.   Asymptotic Properties of Covariance Estimates

In the following sections, the asymptotic properties of covariance estimates are systematically examined. The results derived below demonstrate that spatial smoothing is not needed in order to locate incoherent signals,

but a certain minimal amount appears to be necessary for accurate direction finding in a specular multipath environment.

Under the usual ergodic assumption, the asymptotic properties of certain useful covariance estimates can be determined from the theoretical results presented in Section B.1 of this chapter. Under very general conditions, the maximum number of incoherent signals that can be theoretically resolved (without error) is one less than the number of array elements. We also show that the usual adaptive array approach (i.e., time-averaging) performs no better than standard beam scanning techniques when the received signals are coherent (e.g., specular multipath)

Fortunately, the coherent case can often be reduced to a (partially) incoherent case by appropriately processing the (original) covariance estimate. In Section B.2, we examine the effects of certain spatial averaging operations. In particular, we consider spatial smoothing and forward/backward averaging. These operations are relatively easy to understand when considered separately, but a definitive statement of their combined effects remains somewhat elusive.

1. Ensemble Averaging

In this sub-section, we first examine the properties of the inverse covariance $\underline{R}^{-1}$ in the asymptotic limit of extremely high signal-to-noise ratios. When the signal-in-space covariance $\underline{P}$ is non-singular, it is shown that $\underline{R}^{-1}$ asymptotically projects any vector into the (noise) space orthogonal to the (signal) space spanned by the I ($<$ L) columns of $\underline{V}$ ($\underline{\alpha}$). An immediate consequence is that $\underline{R}^{-1}$ asymptotically "nulls" only the true direction vectors, i.e.,

$$\underline{P}^{-1} \text{ exists} \Rightarrow \lim_{\text{SNR}\to\infty} \; \underline{v}(\alpha) = \underline{0} \text{ iff } \alpha \in \underline{\alpha} \quad . \tag{2.31}$$

This fundamental result suggests that, for sufficiently large SNR, the "null spectrum"

$$T(\alpha) \triangleq \frac{\underline{v}^H(\alpha)\, \underline{R}^{-1} \underline{v}\,(\alpha)}{|\underline{v}(\alpha)|^2} \tag{2.32}$$

should reveal the true location of the signals by its zeroes. Of course, (2.31) is not applicable in the <u>coherent</u> case because the signal-in-space covariance $\underline{P}$ is singular. In fact, it is easy to demonstrate that the null spectrum idea does not generally work for <u>coherent</u> signals. Given any coherent signal $\underline{s}$ in spatially white noise, we may set $\underline{S} = \underline{s}\,\underline{s}^H$ and $\underline{N} = \underline{I}$. The inverse covariance for this case is easy to compute, i.e.,

$$\underline{R}^{-1} = (\underline{S} + \underline{N})^{-1}$$

$$= (\underline{I} + \underline{s}\,\underline{s}^H)^{-1}$$

$$= \underline{I} - \frac{1}{1 + |\underline{s}|^2}\, \underline{s}\,\underline{s}^H \quad .$$

Consequently, for the coherent case we get

$$T(\alpha) = 1 - \frac{|\underline{s}|^2 |\underline{v}^H(\alpha)\,\underline{s}|^2}{1 + |\underline{s}|^2 |\underline{v}(\alpha)|^2 |\underline{s}|^2} \quad . \tag{2.33}$$

A simple application of the Schwartz inequality reveals that this null spectrum has no zeroes except in the (asymptotic) case of a single dominant signal, i.e.,

$$\lim_{|\underline{s}|^2 \to \infty} T(\alpha) = 0 \qquad \text{iff } \underline{s}' = \underline{v}'(\alpha)$$

where $\underline{s}'$ and $\underline{v}'(\alpha)$ are the appropriately normalized versions of $\underline{s}$ and $\underline{v}(\alpha)$. Even the minima of the null spectrum are no help, since they always occur at the maxima of the signal's "conventional" power spectrum,

$$\Psi(\alpha) \triangleq |\underline{v}'^H(\alpha)\,\underline{s}'|^2 \quad .$$

The relationship between T and Ψ follows easily from (2.33), i.e.,

$$T(\alpha) = 1 - \rho \ \psi(\alpha) \quad ; \qquad \rho \ \overset{\Delta}{=} \ \frac{|\underline{s}|^2}{1 + |\underline{s}|^2} \quad .$$

In light of the preceding discussion, the ergodic property of sample covariance matrices strongly suggests than a non-singular signal-in-space covariance is the critical ingredient in any direction-finding recipe based on standard adaptive array theory. On the other hand, the estimate derived from the "empirical" least-squares covariance method converges to a spatially smoothed version of the true covariance, i.e.,

$$\lim_{K \to \infty} \hat{\underline{R}}_M (K) = \underline{R}_M \quad .$$

This point is stressed because the following argument is presented in terms of an unsmoothed covariance. The beneficial effect of spatial smoothing is explored in sub-section B.2 below.

Before proceeding with the proof of (2.31), we wish to emphasize that this result is valid in a far more general context than the special case of an ideal linear array. In fact, the only property of direction vectors used in the proof is the linear independence of the columns if $\underline{V}$. Since we (must) assume that $\underline{V}$ has at least as many rows as columns, the linear dependence hypothesis guarantees that

$$\underline{W} \overset{\Delta}{=} \underline{V}^H \underline{V}$$

always has an inverse. In turn, $\underline{V}$ will always have a (left) pseudo-inverse,

$$\underline{V}^+ \overset{\Delta}{=} \underline{W}^{-1} \underline{V}^H \quad . \tag{2.34}$$

Clearly, $\underline{V}^+ \underline{V} = \underline{I}$ but $\underline{V} \ \underline{V}^+ \neq \underline{I}$ unless $\underline{V}$ is square. In general, $\underline{V} \ \underline{V}^+$ is "merely" a projection matrix. For example, $(\underline{V} \ \underline{V}^+)^2 = \underline{V} \ \underline{V}^+$ follows easily from the fundamental property of a (left) pseudo-inverse, and $\underline{V} \ \underline{V}^+$ inherits its Hermitian property from $\underline{W}$, i.e.,

$$(\underline{V}\ \underline{V}^+)^H = \underline{V}^{+H}\underline{V}^H$$

$$= \underline{V}\ \underline{W}^{-1}\ \underline{V}^H$$

$$= \underline{V}\ \underline{V}^+ \quad .$$

Naturally, $\underline{V}\ \underline{V}^+\underline{x}$ projects an arbitrary vector $\underline{x}$ into the space spanned by the columns of $\underline{V}$. Since this space "contains" the received signal $\underline{s} = \underline{V}\ \underline{p}$ (e.g., $\underline{V}\ \underline{V}^+\underline{s} = \underline{V}\ \underline{V}^+\underline{V}\underline{p} = \underline{V}\underline{p} = \underline{s}$), we will refer to the space spanned by the columns of $\underline{V}$ as the <u>signal space</u>. Obviously, any vector can be written as the sum of its "signal" component and an "orthogonal" component, e.g.,

$$\underline{x} = \underline{V}\ \underline{V}^+\underline{x} + (\underline{I} - \underline{V}\ \underline{V}^+)\ \underline{x} \quad . \tag{2.35}$$

Thus, even a noise vector $\underline{n}$ has a "signal" component that, in effect, causes the <u>signals-in-space</u> to "scintillate" but does not disturb the directional information. Consequently, the complementary space orthogonal to the signal space can justifiably be called the <u>noise</u> space. Of course, the projection matrix for the noise space is $\underline{I} - \underline{V}\ \underline{V}^+$.

Assuming spatially white noise with a power level of unity, the ensemble covariance matrix of the observations (snapshots) follows directly from (2.16) and (2.17)

$$\underline{R} = \underline{V}\ \underline{P}\ \underline{V}^H + \underline{I} \tag{2.36}$$

The inverse of $\underline{R}$ can be written as

$$\underline{R}^{-1} = \underline{I} - \underline{V}\ \underline{Q}\ \underline{V}^H \tag{2.37}$$

where the Hermitian matrix $\underline{Q}$ satisfies

$$\underline{P} - \underline{Q} = \underline{P}\ \underline{W}\ \underline{Q} \quad . \tag{2.38}$$

Since $(\underline{I} + \underline{P}\ \underline{W})^{-1}$ always exists, the unique solution to (2.38) is

$$\underline{Q} = (\underline{I} + \underline{P}\ \underline{W})^{-1}\underline{P} \quad . \tag{2.39}$$

When $\underline{P}$ is non-singular, we also have

$$\underline{Q} = (\underline{P}^{-1} + \underline{W})^{-1} \quad . \tag{2.40}$$

If we take the signal-to-noise ratio (SNR) to be the smallest eigenvalue of $\underline{P}$, it follows that

$$\lim_{SNR \to \infty} \underline{Q} = \underline{W}^{-1} \quad . \tag{2.41}$$

Using this result in (2.37) gives

$$\lim_{SNR \to \infty} \underline{R}^{-1} = \underline{I} - \underline{V}\ \underline{W}^{-1}\underline{V}^H \quad . \tag{2.42}$$

$$= \underline{I} - \underline{V}\ \underline{V}^+$$

Therefore, $\underline{R}^{-1}$ asymptotically approaches the noise-space projection matrix, and (2.31) follows trivially, e.g.,

$$\lim_{SNR \to \infty} \underline{R}^{-1}\underline{V} = (\underline{I} - \underline{V}\ \underline{V}^+)\underline{V} \quad . \tag{2.43}$$

$$= \underline{0}$$

Conversely, suppose $\underline{v}(\alpha')$ is a direction vector such that

$$\lim_{SNR \to \infty} \underline{R}^{-1}\ \underline{v}(\alpha') = \underline{0}$$

it then follows from (2.42) and the standard decomposition in (2.35) that

$$\underline{v}(\alpha') = \underline{V}\ \underline{V}^+\underline{v}(\alpha')$$

is a linear combination of the columns of $\underline{V}$. If we assume

$$\alpha' \neq \alpha_i, \; i \in \{1, \ldots I < L\} \qquad ,$$

then $\underline{V}' = \left(\underline{v}(\alpha') | \underline{V} \right)$ has linearly independent columns (by hypothesis), and an immediate contradiction arises. Therefore, we must have

$$\alpha' \in \{\alpha_i | i = 1, \ldots, I < L\} \qquad .$$

Of course, the case $I = L$ is totally uninteresting, since

$$\underline{V}^+ = \underline{V}^{-1} \Rightarrow \underline{I} - \underline{V}\,\underline{V}^+ = \underline{0} \qquad .$$

Although the orthogonality property (2.31) is an encouraging theoretical result, a pragmatist might find the SNR requirement somewhat disconcerting. Evidently, this difficulty could be overcome by working directly with the signal covariance except for the fact that $\underline{S}$ is singular in all cases of interest. However, by proceeding along the same lines as above, it is easy to show that

$$\lim_{\varepsilon \to 0} \varepsilon(\underline{S} + \varepsilon\underline{I})^{-1} = \underline{I} - \underline{V}\,\underline{V}^+ \qquad (2.44)$$

for any SNR, subject to the same conditions required for (2.31). Presumably, this result is too optimistic since it suggests that perfect performance is possible, even at an arbitrarily small SNR, provided enough snapshots (data) are available. Nonetheless, (2.44) naturally leads one to investigate the potential benefits of noise power cancellation (NPC).

Principal Component Anlaysis

In its simplest form, NPC is implemented by subtracting an estimate of the noise level from the diagonal terms of the (estimated) snapshot covariance [51]. Unfortunately, this rudimentary approach does not guarantee a positive-definite matrix. To overcome this technical difficulty, it is suggested that

the noise level estimate for NPC be derived from the minimum eigenvalue of the covariance estimate. This approach is closely related to the method of Pisarenko discussed in section F of Chapter IV.

Finding the minimum eigenvalue of a covariance matrix is a non-trivial task, especially when other roots of the characteristic equation

$$\underline{R}\,\underline{e} = \lambda\underline{e} \tag{2.45}$$

have nearly the same value. Algorithms are available that solve (2.45), but the eigenvalue approach to NPC may be ruled out by computational cost/complexity considerations. If not, several other super-resolution techniques based on eigenalysis deserve at least as much attention.

It follows immediately from the characterisitc equation (2.45) that the eigenvectors are unaffected by any "adjustment" applied (uniformly) to the main diagonal of $\underline{R}$. Consequently, schemes based on the eigenvectors should be insensitive to any uncertainty in the "correct" noise level.

Here, we use the term "principal components" to mean the orthonormalized eigenvectors of the positive eigenvalues of a covariance matrix. Any (eigen) vector with an eigenvalue of zero is called a null vector. Let $\underline{E}_S$ denote a matrix constructed (column fashion) from the principal components of $\underline{S}$. In similar fashion, a matrix $\underline{E}_N$ can be constructed from any orthogonal basis for the null space (subspace of null vectors) of $\underline{S}$. Since eigenvectors of a Hermitian matrix with distinct eigenvalues are always orthogonal, it follows that

$$\underline{E}_N^H\,\underline{E}_S = \underline{0} \quad . \tag{2.46}$$

Taken together, the principal components of $\underline{E}_S$ and the (orthonormal) null vectors in $\underline{E}_N$ provide a convenient resolution of the identity, i.e.,

$$\underline{E}_S\underline{E}_S^H + \underline{E}_N\underline{E}_N^H = \underline{I} \quad . \tag{2.461}$$

Theorem: The principal components of $\underline{S}$ always span a subspace of the signal space. When $\underline{P}^{-1}$ exists, the principal components span the _entire_ signal space.

To prove the preceding statement (theorem), we first write $\underline{S}$ in terms of the standard spectral (eigen) decomposition for a Hermitian matrix, i.e.,

$$\underline{S} = \sum_{\ell=1}^{L} \lambda_\ell \underline{e}_\ell \underline{e}_\ell^H \qquad (2.47)$$

where $\{\underline{e}_\ell \mid \ell=1, \ldots, L\}$ is a (complete) orthonormal basis consisting of eigenvectors of $\underline{S}$. Combining (2.47) and (2.17), we may write

$$\underline{V}\,\underline{P}\,\underline{V}^H = \underline{E}_S \Lambda_S \underline{E}_S^H \qquad (2.48)$$

where $\underline{\Lambda}_S$ is the diagonal matrix of (positive) eigenvalues of $\underline{S}$ corresponding to its principal components. Post-multiplying both sides of (2.48) by $\underline{E}_S \Lambda_S^{-1}$, it follows immediately that

$$\underline{E}_S = \underline{V}\,\underline{X} \qquad , \qquad (2.49)$$

where

$$\underline{X} = \underline{P}\,\underline{V}^H \underline{E}_S \Lambda_S^{-1} \qquad .$$

This result shows that every principal component of $\underline{S}$ can be written as a linear combination of the true direction vectors. Thus, the principal components of $\underline{S}$ always lie _in_ the signal space.

Provided the signal-in-space covariance is non-singular, next premultiply (2.48) by $\underline{P}^{-1}\underline{V}^+$ to get

$$\underline{V}^H = \underline{P}^{-1}\,\underline{V}^+ \underline{E}_S \Lambda_S \underline{E}_S^H \qquad .$$

Taking the Hermitian transpose leads easily to

$$\underline{V} = \underline{E}_S \underline{Y} \tag{2.50}$$

where

$$\underline{Y} = \underline{\Lambda}_S (\underline{V}^+ \underline{E}_S)^H \underline{P}^{-1} \quad .$$

Comparing (2.49) and (2.50), one concludes that the principal components and the true direction vectors span the same space (since each principal component is a linear combination of the true direction vectors and vice versa). In fact, it follows from (2.49), (2.50) and the definition (2.34) of $\underline{V}^+$ that the projection matrices derived from the principal components and the true direction vectors are the same, e.g.,

$$\underline{V}^+ = (\underline{V}^H \underline{V})^{-1} \underline{V}^H$$

$$= (\underline{Y}^H \underline{E}_S^H \underline{E}_S \underline{Y})^{-1} \underline{Y}^H \underline{E}_S^H$$

where the last step follows from the orthonormality of principal components,
$$= (\underline{Y}^H \underline{Y})^{-1} \underline{Y}^H \underline{E}_S^H$$
i.e.,

$$\underline{E}_S^H \underline{E}_S = \underline{I} \iff \underline{E}_S^H = \underline{E}_S^+ \quad ,$$

therefore,

$$\underline{V}\,\underline{V}^+ = \underline{E}_S \underline{Y} \,(\underline{Y}^H \underline{Y})^{-1} \underline{Y}^H \underline{E}_S^H \quad . \tag{2.51}$$

Substituting (2.49) in (2.50) gives $\underline{V} = \underline{V}\,\underline{X}\,\underline{Y}$. Since $\underline{V}$ has a pseudo-inverse, we get $\underline{I} = \underline{X}\,\underline{Y}$. Similarly, substituting (2.50) in (2.49) leads to $\underline{I} = \underline{Y}\,\underline{X}$. Since $\underline{X}$ is both a left and a right inverse of $\underline{Y}$, $\underline{Y}$ must be square and $\underline{Y}^{-1} = \underline{X}$ exists. Consequently, (2.51) reduces to the desired result, i.e.,

$$\underline{V}\,\underline{V}^+ = \underline{E}_S \underline{E}_S^H \tag{2.52}$$

$$= \underline{E}_S \underline{E}_S^+ \quad .$$

Since the null space is complementary to the space spanned by the principal components of $\underline{S}$, it follows directly from (2.52) and (2.461) that

$$\underline{I} - \underline{V}\,\underline{V}^+ = \underline{E}_N\underline{E}_N^H$$

$$= \underline{E}_N\underline{E}_N^+ \qquad . \tag{2.53}$$

Therefore, under the fundamental assumption that $\underline{P}^{-1}$ exists, the results in (2.42) and (2.53) may be combined to yield

$$\lim_{SNR\to\infty} \underline{R}^{-1} = \underline{E}_N\underline{E}_N^+ \qquad . \tag{2.531}$$

Under the same conditions for (2.31), we also have the equivalent orthogonality property for the "null" space, i.e.,

$$\underline{P}^{-1} \text{ exists} \Rightarrow \underline{E}_N^+\underline{v}(\alpha) = \underline{0} \text{ iff } \alpha \in \underline{\alpha} \tag{2.532}$$

2. Spatial Averaging

In this section, we consider several operations designed to "smooth" sample covariance matrices. We first consider the spatial smoothing operation introduced previously in Section A of this chapter. In the limiting case of perfectly coherent signals, the spatial smoothing operation leads to an effective signal-in-space covariance matrix with rank determined by the number of submatrices averaged and the number of signals present. Subsequently, it is shown that the direction-finding capability of an (ideal) array with L elements is no worse than [L/2] coherent signals.

We next introduce a convenient notational device called the exchange for the purpose of exploiting array symmetry. Consequently, forward/backward averaging emerges as a likely candidate for increasing the coherent signal capability of an array. In favorable situations, the forward/backward averaging operation can effectively increase the rank of the signal-in-space matrix by as much as a factor of two. However, the combined effect of spatial smoothing and forward/backward averaging cannot increase an

array's direction-finding capability beyond [2L/3] coherent signals. Moreover, pathological (e.g., in-phase) situations exist where forward/backward averaging offers no improvement whatsoever.

Besides spatial smoothing and forward/backward averaging, a class of operations exist that lead to Toeplitz covariance (i.e., lag) estimates. Several such methods have been examined and found to perform very poorly in coherent signal environments. In our opinion, the problem of estimating (correlation) lags given a limited amount of data has not yet been satisfactorily[†] solved. For the sake of completeness, the more common lag estimates are briefly reviewed in Chapter IV.

Spatial Smoothing

Since the spatial smoothing operation is linear, its effect on the signal and noise components of the true covariance may be considered separately, i.e.,

$$\underline{R}_M = \underline{S}_M + \underline{N}_M \tag{2.54}$$

Here, we consider only the signal covariance for an ideal linear array. The approach is valid for any signal-in-space, but the results are most important in the coherent case.

We begin by writing the signal received on the mth subaperture as

$$\underline{s}_m = \underline{A}_m \underline{s} = \underline{A}_m \underline{V} \underline{p} \tag{2.55}$$

where $\underline{A}_m$ is a sampling matrix of the form specified previously (2.24). Given

---

[†] For example, given independent observations of a (general) Gaussian vector process, it is well known that the maximum likelihood (ML) covariance estimate is the sample covariance matrix. In contrast, when the true covariance is known to be Toeplitz, none of the standard lag estimation techniques yield the true ML lag estimates.

identical (i.e., ideal) subapertures, each containing N elements, we claim that the subaperture signal(s) can be written as

$$\underline{s}_m \approx \underline{V}_N \underline{p}_{(m)} \tag{2.56}$$

where $\underline{V}_N$ denotes the matrix of true direction vectors for <u>any</u> of the subapertures. Thus, the dependence of $\underline{s}_m$ on the subaperture index has been completely absorbed by the vector $\underline{p}_{(m)}$, which represents the signal-in-space at the phase center of the mth subaperture.

Now consider a <u>hypothetical</u> array with M elements at the subaperture phase centers, distributed symmetrically with respect to the center of the actual array. The "received" signals at the subaperture phase centers are specified by the M-dimensional column vector $\underline{V}_M \underline{p}$. Since the signal received at an element is simply the superposition (sum) of the signals-in-space <u>at that element</u>, $\underline{p}_{(m)}$ must be the Hadamard (term-by-term) product of $\underline{p}$ with the mth row of $\underline{V}_M$. Therefore, the desired signal-in-space vectors can be conveniently represented in matrix form by

$$\left(\underline{p}_{(1)} \cdots \underline{p}_{(M)}\right) = \begin{pmatrix} p_1 & & 0 \\ & \ddots & \\ 0 & & p_J \end{pmatrix} \underline{V}_M^T \tag{2.57}$$

The spatially-smoothed signal covariance "estimate" obtained from (2.56) is

$$\begin{aligned}
\underline{\hat{S}}_M &\triangleq \frac{1}{M} \sum_{m=1}^{M} \underline{s}_m \underline{s}_m^H \\
&= \underline{V}_N \left[\frac{1}{M} \sum_{m=1}^{M} \underline{p}_{(m)} \underline{p}_{(m)}^H\right] \underline{V}_N^H
\end{aligned} \tag{2.58}$$

The term in brackets is

$$\underline{P}_{(M)} \triangleq \frac{1}{M} \left(\underline{p}_{(1)} \cdots \underline{p}_{(M)}\right) \begin{bmatrix} \underline{p}_{(1)}^H \\ \vdots \\ \underline{p}_{(M)}^H \end{bmatrix}$$

$$= \frac{1}{M} \begin{vmatrix} p_1 & \cdot & \cdot & 0 \\ 0 & \cdot & \cdot & p_I \end{vmatrix} \underline{V}_M^T \underline{V}_M^* \begin{vmatrix} p_1^* & \cdot & \cdot & 0 \\ 0 & \cdot & \cdot & p_I^* \end{vmatrix} \cdot \tag{2.59}$$

The determinant of $\hat{\underline{P}}_{(M)}$ follows easily, i.e.,

$$\det \left\{ \hat{\underline{P}}_{(M)} \right\} = \det \left\{ M^{-1} \underline{W}_{(M)}^* \right\} \prod_{i=1}^{I} |p_i|^2 \tag{2.60}$$

where

$$\underline{W}_{(M)} = \underline{V}_M^H \underline{V}_M \tag{2.61}$$

From (2.55), (2.58) and the definition preceding (2.59), we have

$$\underline{S}_M = E\{\hat{\underline{S}}_M\}$$

$$= \underline{V}_N \underline{P}_{(M)} \underline{V}_N^H \; ; \quad N = L + 1 - M \tag{2.62}$$

where

$$\underline{P}_{(M)} \overset{\Delta}{=} E\left\{ \frac{1}{M} \sum_{m=1}^{M} \underline{P}_{(m)} \underline{P}_{(m)}^H \right\}$$

$$= E\{\hat{\underline{P}}_{(M)}\} \qquad \cdot \tag{2.63}$$

In the incoherent (stationary) case, the signal-in-space covariance $\underline{P}$ is a (real) diagonal matrix, say $\underline{\Psi}$. From (2.57), we may write

$$\underline{P}_{(m)} = \underline{D}_{(m)} \underline{p}$$

where $\underline{D}_{(m)}$ is a diagonal matrix constructed from the mth row of $\underline{V}_M$. Since the elements of an _ideal_ direction vector all have unit magnitude, it follows that

$$E\{\underline{P}_{(m)} \underline{P}_{(m)}^H\} = \underline{D}_{(m)} E\{\underline{p} \, \underline{p}^H\} \underline{D}_{(m)}^H \; ; \quad \text{arbitrary } m \qquad \cdot$$

$$= \underline{\Psi}$$

Therefore, nothing is "gained" by spatial smoothing in the incoherent case except a "loss" of aperture, i.e.,

$$\underline{S}_{L+1-N} = \underline{V}_N \, \underline{\Psi} \, \underline{V}_N^H \qquad \text{(incoherent case)} \qquad .$$

On the other hand, a slightly inefficient use of the available aperture may be well worth trading for multipath immunity.

In the coherent case, it suffices to establish conditions that guarantee a non-singular $\hat{\underline{P}}_{(M)}$ in (2.59). From (2.60), we see that this condition is met if $\underline{W}_{(M)}$ is non-singular (we assume that the signals-in-space do not vanish!). It then follows from (2.61) that $\hat{\underline{P}}_{(M)}$ is non-singular if and only if $M \geq I$ (since $\underline{V}_M$ is an MxI matrix of distinct direction vectors). The maximal amount of spatial smoothing is constrained by $M \leq L-I$, and a non-singular $\hat{\underline{P}}_{(M)}$ cannot be achieved by spatial smoothing if $I > L/2$.

Assuming $I < L/2$, the minimum number of sub-apertures M needed to guarantee a non-singular signal-in-space covariance (after spatial smoothing) is $[L/2]$. Each subaperture must contain a minimum of I+1 elements in order to apply the asymptotic results of the previous section to the spatially-averaged (subaperture) covariance. Thus, the number of subaperture elements $N = L + 1 - M$ must be larger than $[L/2]$. These conditions are just met for $I = L/2$, by choosing $M = L/2$ and $N = L/2 + 1$.

We may conclude that perfect direction-finding in a specular multipath environment is theoretically (i.e., asymptotically) possible with a sampled aperture system, provided the number of signals present does not exceed half the number of elements in the array. In the next section, another form of spatial averaging is examined in an attempt to increase the maximum number of (coherent) signals that can be resolved.

Forward/Backward Averaging

Consider the vector $\underline{r}^\dagger$ that results from reversing and conjugating the data samples from an array with L elements, i.e.,

$$r_\ell^\dagger \triangleq r_{L+1-\ell}^* \quad ; \quad \ell = 1, \ldots, L \qquad (2.64)$$

This relationship may be expressed in matrix form

$$\underline{r}^\dagger = \underline{J}\,\underline{r}^* \qquad (2.65)$$

by introducing the (LxL) exchange matrix

$$\underline{J} \triangleq \begin{matrix} & 0 & & 1 \\ & & \cdot & \\ & 1 & & 0 \end{matrix} \qquad (2.66)$$

In general, the exchange of $\underline{A}$ is defined here to be

$$\underline{A}^\dagger \triangleq \underline{J}\,\underline{A}^*\underline{J}' \qquad (2.67)$$

where $\underline{J}$ and $\underline{J}'$ are exchange matrices of the appropriate dimensions. Since every (real) exchange matrix is both symmetric and orthogonal (i.e., $\underline{J}^2 = \underline{I}$), a matrix can only be exchanged once, i.e., for any $\underline{A}$,

$$\underline{A}^{\dagger\dagger} = \underline{A} \qquad (2.68)$$

The exchange operator distributes exactly like the complex conjugate, i.e.,

$$(\underline{A}\,\underline{B})^\dagger = \underline{A}^\dagger\underline{B}^\dagger \quad , \qquad (2.69)$$

and is quasi-linear, i.e.,

$$(\underline{A} + \underline{B})^\dagger = \underline{A}^\dagger + \underline{B}^\dagger \quad , \qquad (2.70a)$$

but

$$(c\underline{A})^\dagger = c^*\underline{A}^\dagger \qquad (2.70b)$$

for any (complex) scalar c. Since $\underline{J} = \underline{J}^* = \underline{J}^T = \underline{J}^{-1}$, it can be shown that the exchange operator commutes with all of the standard superscript operators used here, including the pseudo-inverse.

Comparing (2.03) and (2.64), we deduce the following important property of the direction vectors in an ideal array

$$\underline{v}^\dagger(\alpha) = \underline{v}(\alpha) \quad .\tag{2.71}$$

In fact, (2.71) holds for any linear array of isotropic (i.e., identical) elements distributed **symmetrically** about the array phase center. In general, any matrix (or vector) equal to its exchange will be called harmonic. It follows from (2.69) that the (pseudo-) inverse of any harmonic matrix is also harmonic.

In the remainder of this report, the arithmetic average of $\underline{A}$ and its exchange $\underline{A}^\dagger$ is referred to as the forward/backward average of $\underline{A}$. In order to avoid needlessly cumbersome expressions, let us agree to write the forward/backward average of $\underline{A}$ as

$$_2\underline{A} \stackrel{\Delta}{=} \frac{1}{2}(\underline{A} + \underline{A}^\dagger) \quad .\tag{2.72}$$

Of course, any forward/backward average $_2\underline{A}$ is harmonic.

In general, the covariance of an exchanged vector is the exchange of the covariance of the original vector, i.e,

$$E\{\underline{r}^\dagger \underline{r}^{\dagger H}\} = E\{(\underline{r}\,\underline{r}^H)^\dagger\}$$

$$= \underline{R}^\dagger \quad .\tag{2.73}$$

While it can be shown that any Toeplitz Hermitian matrix (e.g., a correlation matrix) is equal to its exchange, Hermitian matrices are not generally harmonic. A harmonic Hermitian matrix is sometimes called a Hermitian persymmetric matrix [52]. It can also be shown, but is not needed here, that every har-

monic Hermitian matrix has a (complete) orthonormal basis of harmonic eigen-vectors.

Covariance estimates are often forward/backward averaged in order to ensure Hermitian persymmetry. In fact, the forward/backward average of the sample covariance matrix is a sufficient statistic for any stationary Gaussian process with zero mean. Although this statement does not apply in the non-stationary (coherent) case, it has been observed that the performance of the covariance method discussed in Section A.2 generally improves when for-ward/backward averaging is employed. Heuristically, this behavior is ex-plained by the fact that the covariance method is equally applicable to ex-changed array data (because direction vectors are harmonic). An otherwise embarassing dilemma is avoided by simply agreeing to accept the autoregressive coefficients that minimize the average of the prediction errors obtained in both cases. Mathematically, this is easily accomplished by forward/backward averaging the (spatially smoothed) covariance estimate before solving for the autoregressive coefficients. This refinement is referred to here as the modified covariance method, and has previously been studied by several authors [65], [66], [34], [32].

At this point, we wish to demonstrate that forward/backward averaging and spatial smoothing are interchangeable operations (i.e., the order of these operations is immaterial). Because the spatial smoothing operation is linear, it suffices to prove

$$(\underline{R}^{\dagger})_M = (\underline{R}_M)^{\dagger} \quad , \tag{2.74}$$

since it will then follow immediately that

$$\left({}_2\underline{R}\right)_M = \frac{1}{2} \left(\underline{R} + \underline{R}^{\dagger}\right)_M$$

$$= {}_2\left(\underline{R}_M\right) \quad .$$

To establish (2.74), we first write

$$(\underline{R}_M)^\dagger = (\frac{1}{M} \sum_{m=1}^{M} \underline{A}_m \underline{R} \underline{A}_m^T)^\dagger$$

$$= \frac{1}{M} \sum_{m=1}^{M} \underline{A}_m^\dagger \underline{R}^\dagger \underline{A}_m^{\dagger T}$$

(2.75)

by invoking the definition of the spatial smoothing operation (2.25) and the distributive properties of the exchange operator discussed above. Since the (real) sampling matrices that appear in (2.75) all have dimension $N \times 1$ where $N = L + 1 - M$, it is not difficult to show that

$$\underline{A}_m^\dagger = \underline{A}_{M+1-m} \qquad m = 1, \ldots M \quad .$$

(2.76)

Substituting (2.76) into (2.75) and reversing the order of summation yields

$$(\underline{R}_M)^\dagger = \frac{1}{M} \sum_{m=1}^{M} \underline{A}_m \underline{R}^\dagger \underline{A}_m^T$$

$$= (\underline{R}^\dagger)_M$$

as desired.

Our final objective in this section is to determine the effect of forward/backward averaging on any signal covariance matrix of the (general) form $\underline{S} = \underline{V} \underline{P} \underline{V}^H$. First, we find

$$\underline{S}^\dagger = \underline{J}(\underline{V} \underline{P} \underline{V}^H)^* \underline{J}$$

$$(\underline{J} \underline{V}^*)\underline{P}^*(\underline{J} \underline{V}^*)^H$$

$$= \underline{V} \underline{P}^* \underline{V}^H$$

(2.77)

where the last step follows from the harmonic property (2.71) of the columns (direction vectors) of $\underline{V}$. Consequently, forward/backward averaging affects only the signal-in-space covariance, i.e.,

$$
\begin{aligned}
{}_2\underline{S} &= \frac{1}{2}\,(\underline{S} + \underline{S}^\dagger) \\
&= \frac{1}{2}\,\underline{V}\,(\underline{P} + \underline{P}^*)\underline{V}^H \\
&= \underline{V}(\text{Re}\ \underline{P})\underline{V}^H
\end{aligned} \tag{2.78}
$$

where, as usual, Re $\underline{A}$ represents the real part of $\underline{A}$. Of course, the same argument may be applied to the spatially smoothed signal covariance matrix in (2.62) to yield

$$
{}_2\underline{S}_M = \underline{V}_N\big(\text{Re}\ \underline{P}_{(M)}\big)\underline{V}_N^H \quad , \qquad N + M = L + 1 \tag{2.79}
$$

The utility of (2.79) lies in the fact that all of the covariance operations considered here are linear and interchangeable, e.g.,

$$
\begin{aligned}
\lim_{K \to \infty}\ {}_2\hat{\underline{R}}_M(K) &= E\big\{{}_2\hat{\underline{R}}_M(K)\big\} \\
&= {}_2\underline{R}_M \\
&= {}_2\underline{S}_M + {}_2\underline{N}_M \\
&= {}_2\underline{S}_M + \nu\underline{I}
\end{aligned} \tag{2.80}
$$

when the noise is "spatially white."

III. THEORETICALLY OPTIMUM PROCESSORS AND PERFORMANCE

In this chapter, statistical decision theory is applied to the direction-finding problem. In particular, maximum likelihood (ML) angle estimates are discussed in section A, since these estimates are generally asymptotically unbiased and _efficient_. The Cramer-Rao (CR) bound for unbiased angle estimates is derived in section B for the important special case of a single observation. The multiple-snapshot case has also been considered but specific results are not yet available.

Unless otherwise stated, the noise component of the data is assumed here to be a zero-mean, complex Gaussian random vector with uncorrelated, unit variance elements. When necessary, the results are easily extended to the correlated case via the usual "whitening" approach. Two distinct signal models for the received signal have been considered. The easiest way to proceed is to treat the signal deterministically. In this case, all of the uncertainty in the received signal is put in the form of unknown parameters, and as it turns out, the maximum likelihood estimates are relatively easy to specify. Unfortunately, the behavior of the CR bound as the number of snapshots increases is far from transparent. On the other hand, a random (Gaussian) signal model generally contains fewer unknown parameters and leads to CR bounds with a simple inverse (i.e., $1/K$) dependence on the number of snapshots K. The chief drawback of the Gaussian signal model is that the ML estimates and the CR bounds are much more difficult to derive.

A. Maximum Likelihood Estimates

In Section A.1, we first consider the ML estimates of the angles of arrival $\underline{\alpha}$ and the complex signal-in-space amplitudes $\underline{p}$ for the important special case of a single observation of the array elements. Unfortunately, we are unable to obtain a complete solution and must settle for a geometric interpretation of the ML estimate. A gradient projection technique for computing the ML estimate recursively is briefly examined. This particular approach depends on the same information matrix used to compute the Cramer-Rao angle estimation bounds in Section B.

In section A.2 we seek the ML estimate derived from multiple observations, assuming perfectly correlated signals and statistically independent noise. This model is appropriate for coherent specular multipath. Deterministic but unknown amplitude and phase fluctuations for the (common) signal source are postulated. As in the case of a single look, exact formulas are not forthcoming. However, the asymptotic form of the multiple-look ML estimate can be interpreted as a single-look ML estimate derived from the 'principal component' of the observed data. This observation will eventually lead us to introduce an interesting sub-optimal algorithm (EAR) in Chapter IV.

1. Deterministic Signal: One Observation

In the case of a deterministic signal $\underline{s} = \underline{V}(\underline{\alpha})\underline{p}$, the probability density function (pdf) of an arbitrary snapshot $\underline{r}$ is

$$\psi_{\underline{r}} = \frac{1}{(2\pi)^L} \exp\left\{-|\underline{r} - \underline{V}(\underline{\alpha})\underline{p}|^2\right\} \quad . \tag{3.01}$$

The (log) likelihood function is obtained by writing the natural logarithm of the pdf as a function of the unknown parameters. The maximum likelihood estimates are the parameter values that maximize the likelihood function. Thus, the likelihood function obtained from (3.01) may be written as

$$\ell_{\underline{r}}(\underline{\alpha}, \underline{p}) = -|\underline{r} - \underline{V}(\underline{\alpha})\underline{p}|^2 \quad . \tag{3.02}$$

Maximizing (3.02) with respect to $\underline{p}$ is simply another version of the least-squares problem. As usual, the solution is conveniently expressed in terms of the pseudo-inverse operator. Thus, the signal-in-space estimate

$$\hat{\underline{p}}(\underline{\alpha}) \triangleq \underline{V}^+(\underline{\alpha}) \, \underline{r} \tag{3.03}$$

maximizes (3.02) for any given $\underline{\alpha}$. The corresponding estimate of the received signal,

$$\hat{\underline{s}}(\underline{\alpha}) \triangleq \underline{V}(\alpha) \, \hat{\underline{p}}(\underline{\alpha}) \tag{3.04}$$

$$= \underline{V}(\underline{\alpha}) \, \underline{V}^+(\underline{\alpha}) \, \underline{r}$$

is the projection of the observed data, $\underline{r}$, into the space spanned by the (column) direction vectors of $\underline{V}(\underline{\alpha})$. It naturally follows that $\hat{\underline{s}}$ and $\underline{r}-\hat{\underline{s}}$ are orthogonal, i.e.,

$$\hat{\underline{s}}^H (\underline{r} - \hat{\underline{s}}) = 0 \qquad . \tag{3.05}$$

Consequently, the <u>reduced</u> likelihood function takes the following forms:

$$\hat{\ell}_{\underline{r}}(\underline{\alpha}) \quad \overset{\Delta}{=} \max_{\underline{p}} \ell_{\underline{r}}(\underline{\alpha}, \underline{p}) + |\underline{r}|^2$$

$$= |\underline{r}|^2 - |\underline{r} - \hat{\underline{s}}(\underline{\alpha})|^2$$

$$= |\underline{r}|^2 - \underline{r}^H (\underline{r} - \hat{\underline{s}}) + \hat{\underline{s}}^H (\underline{r} - \hat{\underline{s}})$$

$$= \underline{r}^H \hat{\underline{s}}(\underline{\alpha}) \qquad . \tag{3.06}$$

Of course, (3.05) also allows us to write

$$\hat{\ell}_{\underline{r}}(\underline{\alpha}) = |\hat{\underline{s}}(\underline{\alpha})|^2 \qquad . \tag{3.061}$$

Substituting (3.04) now yields

$$\hat{\ell}_{\underline{r}}(\underline{\alpha}) = |\underline{V}(\underline{\alpha})\, \underline{V}^+(\underline{\alpha})\, \underline{r}|^2 \qquad . \tag{3.062}$$

Using the easily verified fact that $\underline{V}\,\underline{V}^+ = \underline{V}^{+H}\underline{W}\,\underline{V}^+$ (recall Eq. 2.34), the reduced likelihood function may also be interpreted in terms of the signal-in-space estimate, i.e.,

$$\underline{r}^H \hat{\underline{s}} = \underline{r}^H \underline{V}\,\underline{V}^+ \underline{r}$$

$$= \underline{r}^H \underline{V}^{+H}\underline{W}\,\underline{V}^+ \underline{r}$$

$$= \hat{\underline{p}}^H \underline{W}\,\hat{\underline{p}} \qquad . \tag{3.07}$$

The exact maximum likelihood solution is conceptually simple, e.g.,

$$\hat{\underline{\alpha}} \;\Longleftarrow\; \max_{\underline{\alpha}} \; \underline{r}^H \hat{\underline{s}} \, (\underline{\alpha}) \qquad , \tag{3.08}$$

but difficult to implement. Even in the special case of one signal, the inner product $\underline{r}^H \underline{v} \, (\alpha)$ must be calculated for all possible angles α of interest. In the case of uniformly spaced elements, a one-dimensional search can be conducted with reasonable efficiency using the fast Fourier transform (FFT). However, a numerical search of $\underline{r}^H \hat{\underline{s}} \, (\underline{\alpha})$ in the multi-dimensional (vector) case would be prohibitively expensive for most real-time applications. Thus, one naturally looks for recursive algorithms that will (hopefully) converge to a (global) maximum of the likelihood function. One such scheme is the projection method described below.

Gradient Projection Method

Given any differentiable complex signal $\underline{s}(\underline{\gamma})$, let $\underline{G}$ denote the matrix derivative of $\underline{s}$ with respect to the parameters $\underline{\gamma}$, i.e.,

$$G_{mn} \;\triangleq\; \frac{\partial s_m}{\partial \gamma_n} \qquad . \tag{3.09}$$

Given an initial estimate $\hat{\underline{\gamma}}$, the new value of the likelihood function caused by a parameter adjustment $\delta\underline{\gamma}$ is

$$\hat{\ell}_{\underline{r}} \, (\delta\underline{\gamma}) \quad \triangleq \quad \ell_{\underline{r}}(\hat{\underline{\gamma}} + \delta\underline{\gamma})$$

$$= \; - \; |\underline{r} - \underline{s}(\hat{\underline{\gamma}} + \delta\underline{\gamma})|^2$$

$$\approx \; - \; |\underline{r} - \hat{\underline{s}} - \hat{\underline{G}}\delta\underline{\gamma}|^2$$

where $\hat{\underline{s}}$ and $\hat{\underline{G}}$ denote the signal and its partial directive (matrix) evaluated at $\hat{\underline{\alpha}}$. Evidently, the likelihood function is incrementally maximized by choosing

$$\hat{\delta\gamma} <= \min_{\delta\gamma} |(\underline{r} - \hat{\underline{s}}) - \hat{\underline{G}}\delta\gamma|^2 \quad .$$

The formal "solution" $\hat{\delta\gamma} \overset{\Delta}{=} \underline{G}^+(\underline{r} - \hat{\underline{s}})$ is usually complex and hence cannot be correct when $\underline{\gamma}$ is a vector of _real_ parameters. In the real case, the correct solution is obtained by constraining the imaginary part of $\underline{\gamma}$ to be zero. This results in

$$\text{Re}\{\hat{\underline{G}}^H\hat{\underline{G}}\} \hat{\delta\gamma} = \text{Re}\{\underline{G}^H(\underline{r} - \hat{\underline{s}})\} \quad .$$

Of course, $\text{Re}\{\hat{\underline{G}}^H\hat{\underline{G}}\}$ must be non-singular to guarantee a (unique) solution. When this condition is not met, the estimation problem is fundamentally unsound, as demonstrated in the next section.

Fisher Information Matrix

The problems of finding the ML estimate and its performance are very closely linked. The crucial step is evaluating and inverting the (Fisher) information matrix. In general, the information matrix $\underline{\Gamma}$ is obtained from the likelihood function via

$$\Gamma_{mn} = -E \left\{ \frac{\partial^2 \ell_{\underline{r}}(\underline{\gamma})}{\partial\gamma_m\partial\gamma_n} \right\} \tag{3.10}$$

where the expectation is with respect to the pdf of the observed vector $\underline{r}$. Substituting $\ell_{\underline{r}}(\underline{\gamma}) = -|\underline{r} - \underline{s}(\underline{\gamma})|^2$ in (3.10) and using the fact that $E\underline{r} = \underline{s}$, it follows that

$$\Gamma_{mn} = -E \left\{ \frac{\partial^2}{\partial\gamma_m\partial\gamma_n} | \underline{r} - \underline{s}(\underline{\gamma})|^2 \right\}$$

$$= 2\text{Re} \left\{ \frac{\partial\underline{s}^H}{\partial\gamma_m} \frac{\partial\underline{s}}{\partial\gamma_n} \right\} \quad . \tag{3.11}$$

Consequently, the information matrix is given by

$$\underline{\Gamma} = 2 \text{ Re } \{\underline{G}^H\underline{G}\} \quad . \tag{3.12}$$

The Cramer-Rao bound on the covariance of **any** unbiased estimate $\hat{\underline{\gamma}}$ is

$$E \{(\hat{\underline{\gamma}} - \underline{\gamma}) \; (\hat{\underline{\gamma}} - \underline{\gamma})^T\} > \underline{\Gamma}^{-1} \quad . \tag{3.13}$$

Thus, a nearly singular information matrix corresponds to very poor estimates of at least some of the parameters. If $\underline{\Gamma}$ is singular, an unbiased estimate of $\underline{\gamma}$ doesn't exist. The Cramer-Rao bound on direction-finding accuracy for a single observation (snapshot) of a deterministic signal in Gaussian noise is derived in Section B of this chapter.

2. Deterministic Signal: Multiple Observations

When multiple snapshots of array data are available, the exact ML receiver becomes even more difficult to implement. However, the form of the estimate suggests an approximation that leads to practical suboptimal processors.

The coherent case is of primary interest here. Unknown fluctuations in the amplitude and phase of the signal **vector** are represented by the unknown complex parameters x(k), i.e.,

$$\underline{r}(k) = x(k)\underline{V}(\underline{\alpha})\underline{p} + \underline{n}(k) \quad , \quad k = 1, \ldots, K \quad . \tag{3.14}$$

Assuming statistically independent observations in (3.14), the likelihood function is obtained by simply adding the individual likelihood functions, i.e.,

$$\ell(\underline{\alpha},\underline{p},\underline{x}) = -\sum_{k=1}^{K} |\underline{r}(k) - x(k)\underline{V}(\underline{\alpha})\underline{p}|^2 \tag{3.15}$$

where $\underline{x}$ is a vector of the unknown complex amplitudes, i.e., $x_k \underline{\Delta} x(k)$. The maximum of (3.15) over the individual components of $\underline{x}$ is easily obtained in terms of the pseudo-inverse of the signal $\underline{s} = \underline{V} \, \underline{p}$, i.e.

$$\hat{\underline{x}}(k) = \underline{s}^+ \underline{r}(k) \quad . \tag{3.16}$$

Consequently,

$$\ell(\underline{a},\underline{p}) \triangleq \max_{\underline{x}} \ell(\underline{a},\underline{p},\underline{x}) + \sum_{k=1}^{K} |\underline{r}(k)|^2$$

$$= -\sum_{k=1}^{K} |\underline{r}(k) - \underline{c} \, \underline{s}^+ \underline{r}(k)|^2 + \sum_{k=1}^{K} |r(k)|^2$$

$$= \sum_{k=1}^{K} |\underline{s} \, \underline{s}^+ \underline{r}(k)|^2$$

$$= \mathrm{tr} \left\{ \underline{s} \, \underline{s}^+ \, \bar{R}(K) \, \underline{s} \, \underline{s}^+ \right\}$$

$$= \frac{\underline{s}^H \hat{R}(K) \underline{s}}{\underline{s}^H \underline{s}}$$

$$= \frac{\underline{p}^H \underline{V}^H \hat{R}(K) \underline{V} \, \underline{p}}{\underline{p}^H \underline{V}^H \underline{V} \, \underline{p}} \quad . \tag{3.17}$$

Maximizing (3.17) over $\underline{p}$ may be accomplished by setting

$$\underline{p} = \underline{V}^+ \underline{y}$$

and maximizing over $\underline{y}$. Therefore,

$$\ell(\underline{a}) \triangleq \max_{\underline{p}} \ell(\underline{a},\underline{p})$$

$$= \max_{\underline{y}} \frac{\underline{y}^H \underline{V}^{+H} \underline{V}^H \hat{R} \, \underline{V} \, \underline{V}^+ \underline{y}}{\underline{y}^H \underline{V}^{+H} \underline{V}^H \underline{V} \, \underline{V}^+ \underline{y}}$$

$$= \max_{\underline{y}} \frac{\underline{y}^H \underline{V} \, \underline{V}^+ \hat{R} \, \underline{V} \, \underline{V}^+ \underline{y}}{\underline{y}^H \underline{V} \, \underline{V}^+ \underline{y}}$$

3-7

$$> \max_{\underline{y}} \frac{\underline{y}^H (\underline{V} \, \underline{V}^+ \hat{\underline{R}} \, \underline{V} \, \underline{V}^+) \underline{y}}{\underline{y}^H \underline{y}} \qquad (3.18)$$

with equality iff $\underline{y} = \underline{V} \, \underline{V}^+ \underline{y}$. Maximizing the expression on the right-hand side of (3.18) is a standard problem in quadratic forms. In general, if $\underline{H}$ is any Hermitian matrix, then

$$\max_{\underline{y}} \frac{\underline{y}^H \underline{H} \, \underline{y}}{\underline{y}^H \underline{y}} = \max \{\lambda \mid \det(\underline{H} - \lambda \underline{I}) = 0\} \triangleq \lambda_{max}\{\underline{H}\} \quad , \qquad (3.19)$$

i.e., one finds the largest eigenvalue λ_{max} of $\underline{H}$. Of course, the maximizing value of $\underline{y}$ is the eigenvector associated with $\lambda_{max}\{\underline{H}\}$.

Theorem:

If $\underline{e}$ is an eigenvector of $\underline{P} \, \underline{H} \, \underline{P}$, where $\underline{P}$ is any idempotent matrix, then $\underline{P} \, \underline{e}$ is also an eigenvector of $\underline{H}$ and has the same eigenvalue as $\underline{e}$.

Proof:

(a)	$\underline{P} \, \underline{H} \, \underline{P} \, \underline{e} = \lambda \underline{e}$;	hypothesis
(b)	$\underline{P}^2 \underline{H} \, \underline{P} \, \underline{e} = \lambda \underline{P} \, \underline{e}$;	premultiplying (a) by $\underline{P}$
(c)	$\underline{P} \, \underline{H} \, \underline{P}^2 \underline{e} = \lambda \underline{P} \, \underline{e}$;	using $\underline{P}^2 = \underline{P}$ in (b)
(d)	$\underline{P} \, \underline{H} \, \underline{P} \, (\underline{P} \, \underline{e}) = \lambda(\underline{P} \, \underline{e})$;	arranging terms in (c), QED.

Consequently, the reduced likelihood function in (3.18) is given by

$$\ell(\underline{\alpha}) = \lambda_{max}\{\underline{V} \, \underline{V}^+ \hat{\underline{R}} \, \underline{V} \, \underline{V}^+\} \qquad (3.19)$$

Thus, the ML receiver for coherent signals finds the projection $\underline{V} \, \underline{V}^+(\alpha)$ that maximizes the largest eigenvalue of the sample covariance (matrix) of the projected data; i.e., using $\hat{\underline{s}} = \underline{V} \, \underline{V}^+ \underline{r}$ as defined previously (3.04), it follows that

$$\hat{\underline{S}}(\alpha) \triangleq \frac{1}{K} \sum_{k=1}^{K} \hat{\underline{s}}(k) \, \hat{\underline{s}}^H(k)$$

$$= \frac{1}{K} \sum_{k=1}^{K} \underline{V} \ \underline{V}^{+} \underline{r}(k) \underline{r}^{H}(k) \underline{V} \ \underline{V}^{+}$$

$$= \underline{V} \ \underline{V}^{+} \hat{\underline{R}} \ \underline{V} \ \underline{V}^{+}$$

As usual, the exact ML estimate is cumbersome to implement. However, in the (coherent case) under consideration, the behavior of the sample covariance matrix $\hat{\underline{R}}$ is governed by a single large eigenvalue, say $\hat{\lambda}$, and its associated eigenvector $\hat{\underline{e}}$, referred to here as the principal component of $\hat{\underline{R}}$. Using the asymptotic approximation

$$\hat{\underline{R}} \simeq \hat{\lambda} \ \hat{\underline{e}} \ \hat{\underline{e}}^{H}$$

in (3.19) leads to

$$\ell(\underline{\alpha}) \simeq \hat{\lambda} \cdot |\underline{V} \ \underline{V}^{+} \hat{\underline{e}}|^{2}$$

In turn, this approximation suggests that an asymptotically efficient estimate can be obtained by treating the principal component of the sample covariance matrix as if it were a single observation of the array data.

B.   Cramer-Rao Resolution Bound

One of the principal difficulties in investigating improved methods for angle determination is the lack of quantitative limits on the achievable performance analogous to the role of channel capacity for communication systems. This bounding problem is difficult in the angle determination case because the processor cannot be assumed to have a priori knowledge as to the number of plane waves which arise. A preliminary study of the system identification issue was made with the results reported in Appendix E. However, the most useful results to date have been obtained for the case where the number of plane waves is known a priori and we are concerned only with the accuracy in estimating the angles of arrivals.

In this section, we derive the Cramer-Rao (CR) bound for any unbiased angle estimate scheme under the following conditions. The received signals are observed in the presence of (spatially) white noise. Only one observation is available. The complex signal amplitudes (in $\underline{p}$) are considered to be unknown parameters.

We first obtain the Fisher information matrix for a single observation of the array signals in white noise. We then calculate the "reduced" information matrix for the angles-of-arrival assuming unknown complex signal amplitudes. Inverting the reduced information matrix yields the CR bound on angular accuracy under very general conditions. Certain algebraic simplifications are then introduced for isotropic arrays.

The general theory is illustrated by two simple examples. In the case of only one signal, we obtain a slight generalization of the angular accuracy formula for an ideal array, derived previously by Manasse [54] and perhaps others. Explicit results for the more interesting case of two signals are presented in the form of an effective degradation (loss) in the available signal-to-noise ratio.

Returning to the problem originally considered in Section III.A, let us agree to write the received signal totally in terms of real parameters, e.g.,

$$\underline{s}(\underline{\alpha}, \underline{\iota}, \underline{\kappa}) = \underline{V}(\underline{\alpha}) \; \underline{p}(\underline{\iota}, \underline{\kappa}) \tag{3.20}$$

where $\underline{\iota}$ and $\underline{\kappa}$ are the (real) in-phase and quadrature components of $\underline{p}$, i.e.,

$$\underline{\iota} + j\underline{\kappa} \stackrel{\Delta}{=} \underline{p} \tag{3.21}$$

The only additional assumption required for the fundamental result in (3.40) is that the vector $\underline{v}(\alpha)$ have a derivative with respect to α, denoted by $\underline{\dot{v}}(\alpha)$. Thus, a small change $\delta\underline{\alpha}$ in the true angles $\underline{\alpha}$ leads to a perturbation in $\underline{V}$ given by

$$\delta\underline{V} \stackrel{\Delta}{=} \underline{V}(\underline{\alpha} + \delta\underline{\alpha}) - \underline{V}(\underline{\alpha})$$

$$= \left(\dot{\underline{v}}(\alpha_1)\delta\alpha_1 \mid \ldots \mid \dot{\underline{v}}(\alpha_I)\delta\alpha_I \right)$$

$$= \dot{\underline{V}}(\underline{\alpha})\left[\delta\underline{\alpha}\right]$$

where $\dot{\underline{V}} = \dot{\underline{V}}(\alpha)$ denotes the matrix of differentiated column vectors, evaluated at the true angles, i.e,

$$\dot{\underline{V}}(\underline{\omega}) \triangleq \left(\dot{\underline{v}}(\alpha_1) \mid \ldots \dot{\underline{v}}(\alpha_I) \right) \qquad (3.22)$$

and the notation $[\underline{x}]$ is used here to represent the diagonal matrix constructed from the elements of an arbitrary vector $\underline{x}$, i.e.,

$$[\underline{x}] \triangleq \begin{bmatrix} x_1 & & 0 \\ & \ddots & \\ 0 & & x_N \end{bmatrix} \qquad (3.23)$$

Thus, small changes in all the parameters lead to a signal increment of the form

$$\delta\underline{s} = (\delta\underline{V})\underline{p} + \underline{V}\delta\underline{p}$$

$$= \dot{\underline{V}} \, [\delta\underline{\alpha}] \, \underline{p} + \underline{V}(\delta\underline{1} + j\delta\underline{\kappa})$$

$$= \dot{\underline{V}} \, [\underline{p}]\delta\underline{\alpha} + \underline{V}\delta\underline{1} + j\underline{V}\delta\underline{\kappa}$$

The partial derivatives of $\underline{s}$ with respect to the signal parameters follow by inspection, i.e.,

$$\underline{G} = \left[\frac{\partial\underline{s}}{\partial\underline{\alpha}} \mid \frac{\partial\underline{s}}{\partial\underline{1}} \mid \frac{\partial\underline{s}}{\partial\underline{\kappa}} \right]$$

$$= \left(\dot{\underline{V}} \, [\underline{p}] \mid \underline{V} \mid j\underline{V} \right) \qquad (3.24)$$

The Fisher information matrix (3.12) follows directly, e.g,

$$\underline{\Gamma} = 2\text{Re} \begin{bmatrix} [\underline{p}^*] \, \underline{\ddot{w}} \, [\underline{p}] & [\underline{p}^*] \, \underline{\dot{w}}^H & j[\underline{p}^*]\underline{\dot{w}}^H \\ \underline{\dot{w}}^*[\underline{p}] & \underline{w} & j\,\underline{w} \\ -j\,\underline{\dot{w}}\,[\underline{p}] & -j\,\underline{w} & \underline{w} \end{bmatrix} \qquad (3.25)$$

where, primarily for convenience, we have introduced

$$\underline{\ddot{w}} \;\triangleq\; \underline{\dot{v}}^H \, \underline{\dot{v}} \;, \qquad (3.26a)$$

$$\underline{\dot{w}} \;\triangleq\; \underline{v}^H \, \underline{\dot{v}} \;, \qquad (3.26b)$$

and, as usual,

$$\underline{w} \;=\; \underline{v}^H \, \underline{v}$$

Arranging the signal-in-space components in a (nuisance) vector $\underline{\nu}$,

$$\underline{\nu}^T \;\triangleq\; (\underline{\iota}^T \mid \underline{\kappa}^T) \;, \qquad (3.27)$$

the information matrix may be partitioned according to

$$\underline{\Gamma} \;=\; \begin{bmatrix} \underline{\Gamma}_{\alpha\alpha} & \underline{\Gamma}_{\alpha\nu} \\ \hline \underline{\Gamma}_{\nu\alpha} & \underline{\Gamma}_{\nu\nu} \end{bmatrix} \qquad (3.28)$$

and its inverse according to

$$\underline{\Gamma}^{-1} = \left[\begin{array}{c|c} \underline{\Gamma}^{\alpha\alpha} & \underline{\Gamma}^{\alpha\nu} \\ \hline \underline{\Gamma}^{\nu\alpha} & \underline{\Gamma}^{\nu\nu} \end{array} \right] \tag{3.29}$$

The submatrices (in 3.28) and (3.29) are related by the Schur's identity [55]. For our purposes here, we need only the fact that

$$\underline{\Gamma}^{\alpha\alpha} = \left(\underline{\Gamma}_{\alpha\alpha} - \underline{\Gamma}_{\alpha\nu} \, \underline{\Gamma}_{\nu\nu}^{-1} \, \underline{\Gamma}_{\nu\alpha} \right)^{-1} \tag{3.30}$$

Comparing (3.25) and (3.28), we may identify

$$\underline{\Gamma}_{\nu\nu} = 2 \left[\begin{array}{c|c} \text{Re } \underline{W} & -\text{Im } \underline{W} \\ \hline \text{Im } \underline{W} & \text{Re } \underline{W} \end{array} \right] \tag{3.31}$$

and

$$\underline{\Gamma}_{\nu\alpha} = 2 \left[\begin{array}{c} \text{Re } \underline{Z} \\ \hline \text{Im } \underline{Z} \end{array} \right] \tag{3.32}$$

where it is expedient to use

$$\underline{Z} \overset{\Delta}{=} \underline{\mathring{W}} \, [\underline{p}] \tag{3.33}$$

temporarily. Of course, we also have

$$\underline{\Gamma}_{\alpha\nu} = 2 \left(\text{Re } \underline{Z}^H \; \middle| \; - \text{Im } \underline{Z}^H \right) \tag{3.34}$$

$$= 2 \left(\text{Re } \underline{Z}^T \; \middle| \; \text{Im } \underline{Z}^T \right)$$

Due to its special structure, $\underline{\Gamma}_{\nu\nu}$ is easily inverted, i.e.,

$$\underline{\Gamma}_{\nu\nu}^{-1} = \frac{1}{2} \left(\begin{array}{c|c} \text{Re } \underline{W}^{-i} & - \text{Im } \underline{W}^{-1} \\ \hline \text{Im } \underline{W}^{-1} & \text{Re } \underline{W}^{-1} \end{array} \right) \tag{3.35}$$

It follows from (3.32) – (3.35) that

$$\frac{1}{2} \underline{\Gamma}_{\alpha\nu} \underline{\Gamma}_{\nu\nu}^{-1} \underline{\Gamma}_{\nu\alpha} = \text{Re}\{\underline{z}^H \underline{W}^{-1} \underline{z}\} \tag{3.36}$$

$$= \text{Re}\{[\underline{p}^*] \; \dot{\underline{W}}^H \underline{W}^{-1} \dot{\underline{W}} [\underline{p}]\}$$

It then follows that the 'reduced' information matrix for $\underline{\alpha}$ is

$$\underline{\Gamma}_{\alpha\alpha}^{(-)} \triangleq \underline{\Gamma}_{\alpha\alpha} - \underline{\Gamma}_{\alpha\nu} \underline{\Gamma}_{\nu\nu}^{-1} \underline{\Gamma}_{\nu\alpha} \tag{3.37}$$

$$= 2 \, \text{Re}\{[\underline{p}^*] \; (\ddot{\underline{W}} - \dot{\underline{W}}^H \underline{W}^{-1} \dot{\underline{W}}) \; [\underline{p}]\}$$

The array factor

$$\ddot{\underline{W}}_\perp \triangleq \ddot{\underline{W}} - \dot{\underline{W}}^H \underline{W}^{-1} \dot{\underline{W}}$$

$$= \dot{\underline{V}}^H \dot{\underline{V}} - \dot{\underline{V}}^H \underline{V} \underline{W}^{-1} \underline{V}^H \dot{\underline{V}}$$

$$= \dot{\underline{V}}^H \; (\underline{I} - \underline{V} \, \underline{W}^{-1} \, \underline{V}^H) \; \dot{\underline{V}}$$

$$= \dot{\underline{V}}^H \; (\underline{I} - \underline{V} \, \underline{V}^+) \; \dot{\underline{V}} \tag{3.38}$$

depends only on the angle $\underline{\alpha}$ through $\underline{V}$ and its 'derivative' $(\dot{\underline{V}})$. If the projection of $\dot{\underline{V}}$ into the (noise) space orthogonal to the (signal) space spanned by the columns of $\underline{V}$ is written as

$$\dot{\underline{V}}_\perp \triangleq (\underline{I} - \underline{V} \, \underline{V}^+) \; \dot{\underline{V}}$$

It follows easily that

$$\ddot{\underline{W}}_\perp = \dot{\underline{V}}_\perp^H \dot{\underline{V}}_\perp$$

is at least positive semi-definite. Moreover, $\ddot{\underline{W}}_\perp$ is non-singular if and only if the columns of $(\dot{\underline{V}}|\underline{V})$ are linearly independent. In turn, the reduced information matrix $\underline{\Gamma}_{\alpha\alpha}^{(-)}$ is guaranteed to have an inverse whenever the array factor $\ddot{\underline{W}}_\perp$ is non-singular. To see this, first express $\underline{p}$ in terms of its amplitude $\underline{\rho}$ and phase $\underline{\phi}$ as the Hadamard product

$$\underline{p} = \underline{\rho} \times e^{j\underline{\phi}} \tag{3.39}$$

where $e^{j\underline{\phi}}$ is an obvious notation for the vector with elements $e^{j\phi}\mathbf{1}$. Substituting (3.39) and the array factor (3.38) in (3.37), the reduced information matrix can be put in a very convenient form, i.e.,

$$\underline{\Gamma}_{\alpha\alpha}^{(-)} = 2[\underline{\rho}] \ \mathrm{Re}\{[e^{-j\underline{\phi}}] \ \ddot{\underline{W}}_\perp \ (\underline{\alpha})[e^{j\underline{\phi}}]\} \ [\underline{\rho}] \tag{3.40}$$

Thus, a non-singular array factor guarantees the existence of a finite CR (angle) resolution bound. It can be easily shown (see Chapter II) that the array factor is real for uniform arrays with ideal (isotropic) elements. Consequently, when the signals-in-space are all in-phase at the (phase) center of an ideal array, the reduced information matrix

$$\underline{\Gamma}_{\alpha\alpha}^{(-)} = 2[\underline{\rho}] \ \ddot{\underline{W}}_\perp \ [\underline{\rho}] \qquad \text{(worst case)}$$

will not have an inverse if the array factor is singular. The implication is that extremely poor performance can be anticipated when more than $L/2$ (in-phase) signals are present.

Isotropic Arrays

In general, the elements of the array factor are obtained from the (Hermitian) array function

$$w(\alpha,\alpha') \overset{\Delta}{=} \underline{v}^H(\alpha') \ \underline{v}(\alpha) \tag{3.41}$$

$$= w^*(\alpha', \ \alpha)$$

For arrays with identical elements, we have

$$w(\alpha,\alpha') \equiv w(\alpha-\alpha',0) \qquad \text{(isotropic assumption)}$$

In this case, it is convenient to introduce the isotropic array function

$$w(\Delta\alpha) \stackrel{\Delta}{=} w(\Delta\alpha,0) \qquad\qquad (3.42)$$

Several general properties of isotropic array functions follow immediately, e.g.,

$$
\begin{aligned}
w^*(\Delta\alpha) &= w^*(\Delta\alpha,0) \\
&= w(0,\Delta\alpha) \\
&= w(-\Delta\alpha,0) \\
&= w(-\Delta\alpha)
\end{aligned}
\qquad\qquad (3.44)
$$

Since

$$\frac{\partial w}{\partial \alpha}(\alpha,\alpha') = \underline{v}^H(\alpha')\,\underline{\dot{v}}(\alpha) \quad , \qquad\qquad (3.45)$$

the elements of $\underline{\dot{W}}$ in the isotropic case are found by evaluating

$$
\begin{aligned}
\dot{w}(\Delta\alpha) &= \frac{\partial}{\partial \alpha} w(\alpha,\alpha') \;\Big|_{\;\alpha-\alpha'\;=\;\Delta\alpha} \\
&= \frac{\partial}{\partial \alpha} w^*(\alpha'-\alpha,0) \;\Big|_{\;\alpha-\alpha'\;=\;\Delta\alpha} \\
&= -\dot{w}^*(-\Delta\alpha)
\end{aligned}
\qquad\qquad (3.46)
$$

at the appropriate separation angles. Similarly, the elements of $\underline{\ddot{W}}$ are generally obtained from

$$\frac{\partial^2 w}{\partial \alpha \partial \alpha'}(\alpha,\alpha') = \underline{\dot{v}}^H(\alpha')\,\underline{\dot{v}}(\alpha) \qquad , \qquad\qquad (3.47)$$

3-16

but in the isotropic case we may evaluate

$$-\ddot{w}(\Delta\alpha) = \frac{\partial^2 w}{\partial\alpha\partial\alpha'} \; (\alpha,\alpha') \; \big|_{\alpha-\alpha' = \Delta\alpha}$$

$$= \frac{\partial^2 w^*}{\partial\alpha\partial\alpha'} \; (\alpha'-\alpha,0) \; \big|_{\alpha-\alpha' = \Delta\alpha}$$

$$= -\ddot{w}^*(-\Delta\alpha) \tag{3.48}$$

instead.   In the isotropic case, one may also assume the following:

$$w(0) = 1 \;\; \Rightarrow \;\; |\underline{v}(\alpha)|^2 \equiv 1 \quad \text{(normalized array gain)}$$

$$\dot{w}(0) = 0 \;\; \Rightarrow \;\; \underline{v}^H(\alpha) \; \dot{\underline{v}}(\alpha) \equiv 0 \quad \text{(defines the array phase center)}$$

$$\xi^2 \; \overset{\Delta}{=} \; -\ddot{w}(0) \;\; \Rightarrow \;\; |\dot{\underline{v}}(\alpha)|^2 \equiv \xi^2 \; \text{(rms array "length")} \quad .$$

For an ideal array (2.03), the phase center is at the geometric center of the array, and the mean square length of the array in terms of the number of antenna elements is

$$\xi^2 = \frac{1}{L} \sum_{\ell=1}^{L} \; (\ell - \frac{L+1}{2})^2$$

$$= (L^2-1)/12 \qquad \text{(ideal array)} \quad .$$

Example: One signal

In this rather trivial case, the array factor reduces to a scalar, i.e.,

$$\ddot{\underline{W}}_1 \; \to \; \xi^2 \; ,$$

and the CR bound, obtained by inverting (3.40), becomes

$$\sigma_\alpha^2 \; \overset{\Delta}{=} \; E(\hat{\alpha} - \alpha)^2 \; \geq \; \frac{1}{2\rho^2 \xi^2} \quad .$$

This result may be expressed in terms of the beamwidth α_B (2.07) and the <u>array SNR</u> $\overset{\Delta}{=} \rho^2$, i.e.,

$$\sigma_\alpha > \frac{\sqrt{3}}{\pi} \cdot \frac{\alpha_B}{\sqrt{2 \times SNR}} \cdot \frac{1}{\sqrt{1 - 1/L^2}} \qquad .$$

<u>Example: Two Signals</u>

In the more interesting case of two signals, let us agree to write

$$\Delta\alpha \overset{\Delta}{=} \alpha_2 - \alpha_1$$

$$w \overset{\Delta}{=} w(\Delta\alpha)$$

$$\dot{w} \overset{\Delta}{=} \dot{w}(\Delta\alpha)$$

$$\ddot{w} \overset{\Delta}{=} \ddot{w}(\Delta\alpha)$$

and

$$\Delta\phi \overset{\Delta}{=} \phi_2 - \phi_1$$

The matrices needed to calculate the array factor are

$$\underline{W} = \begin{bmatrix} 1 & w \\ w^\star & 1 \end{bmatrix}$$

$$\underline{\dot{W}} = \begin{bmatrix} 0 & \dot{w} \\ -\dot{w}^\star & 0 \end{bmatrix}$$

and

$$\underline{\ddot{W}} = \begin{bmatrix} \xi^2 & -\ddot{w} \\ -\ddot{w}^\star & \xi^2 \end{bmatrix}$$

Consequently, the array factor may be written as

$$\underline{\ddot{W}}_1 = \eta(\Delta\alpha)\xi^2 \begin{bmatrix} 1 & -f(\Delta\alpha) \\ -f^\star(\Delta\alpha) & 1 \end{bmatrix}$$

where the array efficiency factor $\eta(\Delta\alpha)$ is given by

$$\eta \triangleq 1 - \frac{|\dot{w}|^2/\xi^2}{1-|w|^2}$$

and the array coherency factor $f(\Delta\alpha)$ is conveniently defined in terms of η, i.e.,

$$f \triangleq \frac{1}{\xi^2\eta} \left(\ddot{w} + \frac{w^* \dot{w}^2}{1 - |w|^2} \right)$$

The CR bound on the absolute accuracy of the angle estimates may be expressed as

$$\sigma_\alpha^2 \geq \frac{1}{2\eta \cdot SNR \cdot \xi^2} \cdot \frac{1}{1 - |f|^2 \cos^2\Delta\beta};$$

where $\Delta\beta \triangleq \Delta\phi + \arg f(\Delta f)$.

The effective loss of SNR compared to the single signal case is described by the efficiency factor

$$\eta(\Delta\alpha,\Delta\beta) \triangleq \eta(\Delta\alpha) \left(1 - |f(\Delta\alpha)|^2 \cos^2\Delta\beta \right)$$

Thus, the most favorable situation is for signals effectively 90° out-of-phase, i.e.

$$\eta(\Delta\alpha, \pm 90°) = \eta(\Delta\alpha)$$

The worst cases are the (effectively) in-phase or out-of-phase situations, i.e.,

$$\eta(\Delta\alpha,0°) = \eta(\Delta\alpha,180°) = \eta(\Delta\alpha) \left(1 - |f(\Delta\alpha)|^2 \right)$$

The SNR efficiency factor $\eta(\Delta\alpha,\Delta\beta)$ has been plotted in Fig. 3-1 as a function of $\Delta\alpha$ for three different values of $\Delta\beta$. An ideal uniform linear array with nine elements was assumed for the calculations.

The bound on the relative accuracy of the angle estimates does not behave in exactly the same manner as the absolute bound, especially when the signals have nearly the same power, i.e.,

$$\sigma_{\Delta\alpha}^2 = \frac{1}{2\eta \cdot SNR \cdot \xi^2} \cdot \frac{1 + 2\rho|f|\cos \Delta\beta + \rho^2}{1 - |f|^2 \cos^2 \Delta\beta}$$

In this result, the signal-to-noise ratio (SNR) is for the _weaker_ of the targets, and ρ is used to represent the _relative_ amplitude of the two targets, i.e.,

$$\rho \overset{\Delta}{=} \rho_2/\rho_1 \quad \text{(assume } \rho_2 < \rho_1)$$

Several comments are in order regarding the efficiency factor:

(1) the efficiency factor depends only on the relative angular separation and relative phase, but not on the relative signal amplitudes. This suggests that the presence of a second signal which is very much smaller than the larger signal can cause a significant degradation in larger signal angle accuracy if one attempts to estimate the angle of arrival of both signals.

(2) the efficiency at small separation angles varies approximately as θ_{sep}^{-n} where $n = 4$ for inphase conditions and $n = 2$ for the quadrature phase condition. This in turn suggests the minimum separation angle for "useful" angle estimation accuracy will be weakly dependent on the SNR.

(3) the restriction to an unbiased estimator may have significantly affected the bounds. Alternative bounding procedures [82, 83] which do not make this assumption are quite difficult to compute numerically. On the other hand, some of the numerical studies accomplished to date have indicated that the absolute angle accuracy bounds are "reasonable" at high SNR [36, 37, 59].

Issues (1) and (3) above should be the subject of detailed investigation in any follow-on studies.
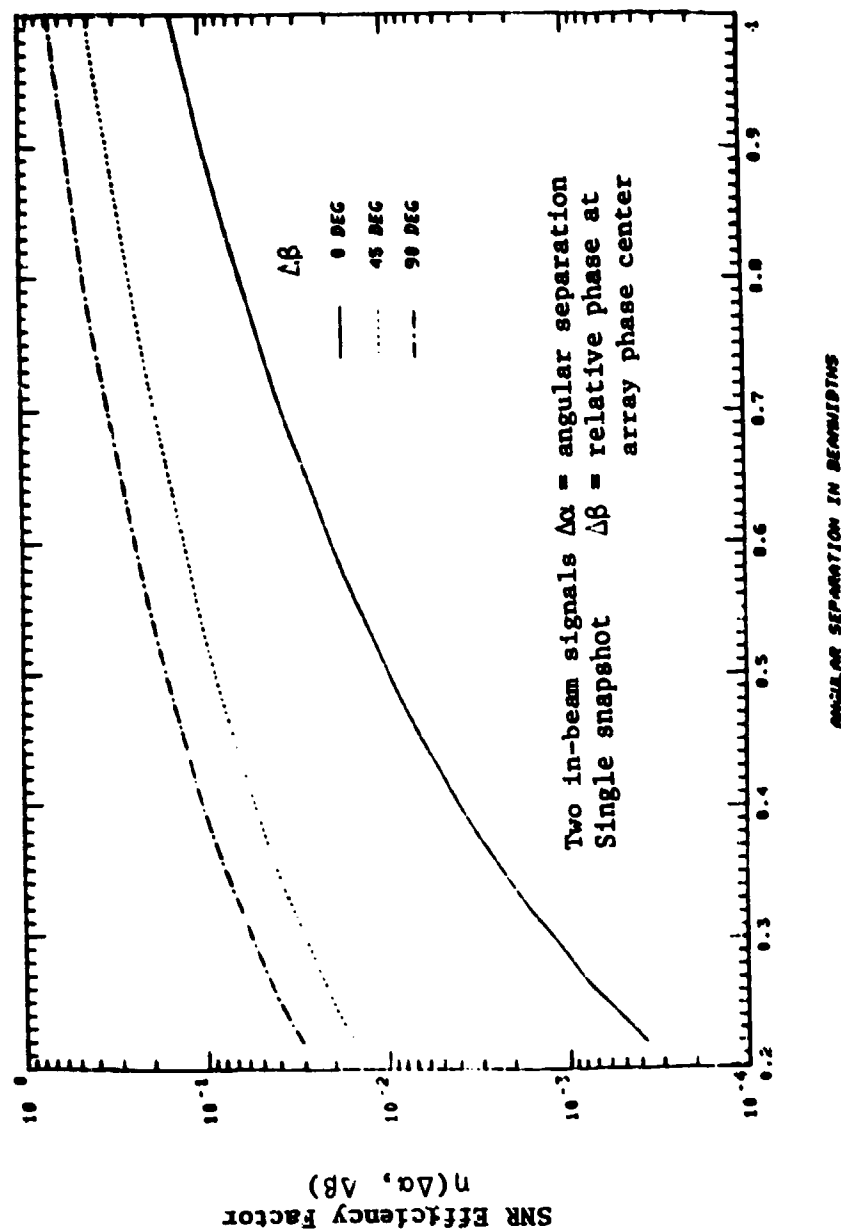
Fig. 3-1. Theoretical degradation in angle estimation accuracy.

IV. CANDIDATE SIGNAL PROCESSING TECHNIQUES

In this chapter, we outline the signal processing techniques investigated during the course of this study. These techniques can be divided into several groups:

(1) approaches discussed by the FRG in the context of DLS and DAS [9,10].

(2) maximum likelihood (ML) parameter estimation methods where one seeks to maximize a probability density function over a specified set of possible signal parameters.

(3) spectral estimation techniques where one first estimates the angular power spectrum and then associates the angle of arrival with the spectrum peaks. The MLS processor for Time Reference Scanning Beam (TRSB) can be viewed as accomplishing this function by more or less standard Fourier methods (see Fig. 4-1). The focus in our studies has been on "high resolution" methods that attempt to distinguish spectral features which cannot be resolved by the standard Fourier techniques.

(4) algebraic methods using eigen expansions of the observed covariance matrix. For a spatially stationary process, the spatial correlation function is the Fourier transform of the angular power spectrum. Consequently, one can analyze the received signal environment in terms of the correlation (or covariance) function.

Section A discusses several of the schemes suggested by the FRG, including interferometer arrays (with and without adaptive element steering), beam multiplexed arrays and noise free multiple signal parameter estimation approaches.

Section B considers the maximum likelihood parameter estimation approach discussed in the preceding chapter. Our investigation consisted primarily of reviewing the previous work in this area in order to assess the feasibility of such an approach for real-time operation.

Adaptive nulling antenna algorithms used in radar systems to mitigate the effects of jammers also have received considerable attention recently. Unfortunately, in the cases of greatest interest here, the standard adaptive

4-1

Fig. 4-1. Relationship of TRSB angle estimation to sampled aperture spectrum analysis.

4-2

nulling algorithms are not applicable because the interference (especially, multipath) cannot be observed in the absence of the desired signal. However, the so-called "maximum liklihood method" (MLM) of spectrum estimation, first proposed by Capon [12], can be applied in these cases and is discussed in Section C.

Maximum entropy (ME)/autoregressive (AR) techniques are considered in Section D. These techniques have considerable practical interest because of their ability to achieve angular accuracies approaching that of the Cramer-Rao bound. Moreover, the mathematical equations used are comparatively easy to implement.* The relationship between various ME/AR interpretations is discussed and the most robust approaches are examined in greater detail.

One interpretation of the ME/AR technique is that received data samples are represented as the output of an all-pole system driven by white noise. The inclusion of receiver (output) noise leads to a system model with zeros as well as poles. The resulting autoregressive moving average (ARMA) model is considered in section E. The results in this case are less complete than those in the preceding sections since the inclusion of zeros significantly complicates the estimation problem.

The final section considers covariance eigenanalysis methods. An underlying motivation for these methods is the notion of noise power cancellation, i.e., attempting to remove bias due to the noise along the diagonal of the covariance matrix. Two new approaches are considered. The multiple signal classification (MUSIC) algorithm suggested by R. Schmidt [30] attempts to utilize orthogonality between "signal" and "noise" eigenvectors to identify the various signal directions. However, this procedure was found to encounter significant difficulties when correlated signals (e.g., multipath) are present. Consequently, an alternative procedure is proposed that uses ME/AR techniques to analyze the principal "signal" eigenvectors.

---

*The ME/AR algorithms are closely related to the speech linear predictive coding (LPC) algorithms which are receiving increased use in speech vocoders.

A.   Advanced Signal Processing Techniques Suggested in the DLS/DAS
      Context

During the DLS/DAS development program, a number of advanced signal
processing methods were suggested as methods of achieving improved perfor-
mance. In this section, we summarize the various approaches and indicate the
expected improvement against multipath.

1.   Interferometric Processing

The bulk of the DLS/DAS processing approaches have involved the use of
multiple baseline interferometers. The theory of such interferometers is well
known [27] and many of the DLS schemes were examined in other Lincoln reports
[28, 29]. Provided that the ambiguities are correctly resolved, the DLS/DAS
linear interferometric arrays will typically yield performance which is very
close to that of a beam steering or monopulse processor using the same element
spacing when the received signal is

a)   a single plane wave in additive white noise, or

b)   two inbeam plane waves plus additive white noise.

The principal performance loss relative to standard processing techniques
(e.g., monopulse) arises from the fact that the inbeam region corresponds to
the beam pattern of the elements used to produce the final interferometer
baseline. The recent DAS proposals (Figs. 4-2 and 4-3) suggest the use of six
(software) phase-steered elements to form a "synthetic element" [33] at each
end of the array. The "inbeam"* region for this particular scheme is roughly
plus or minus one effective beamwidth of the synthetic element about the
direct signal. Since the synthetic element beamwidth is substantially larger
than the classical array beamwidth, the inbeam region for such an array
processor is several times larger than that of the corresponding monopulse
processor. Evidently, the function of the steered elements in Fig. 4-3 is to

---

*Inbeam means the separation between the two plane-wave signals in less than
λ/β (radians) where β/λ is the interferometer baseline in wavelengths.
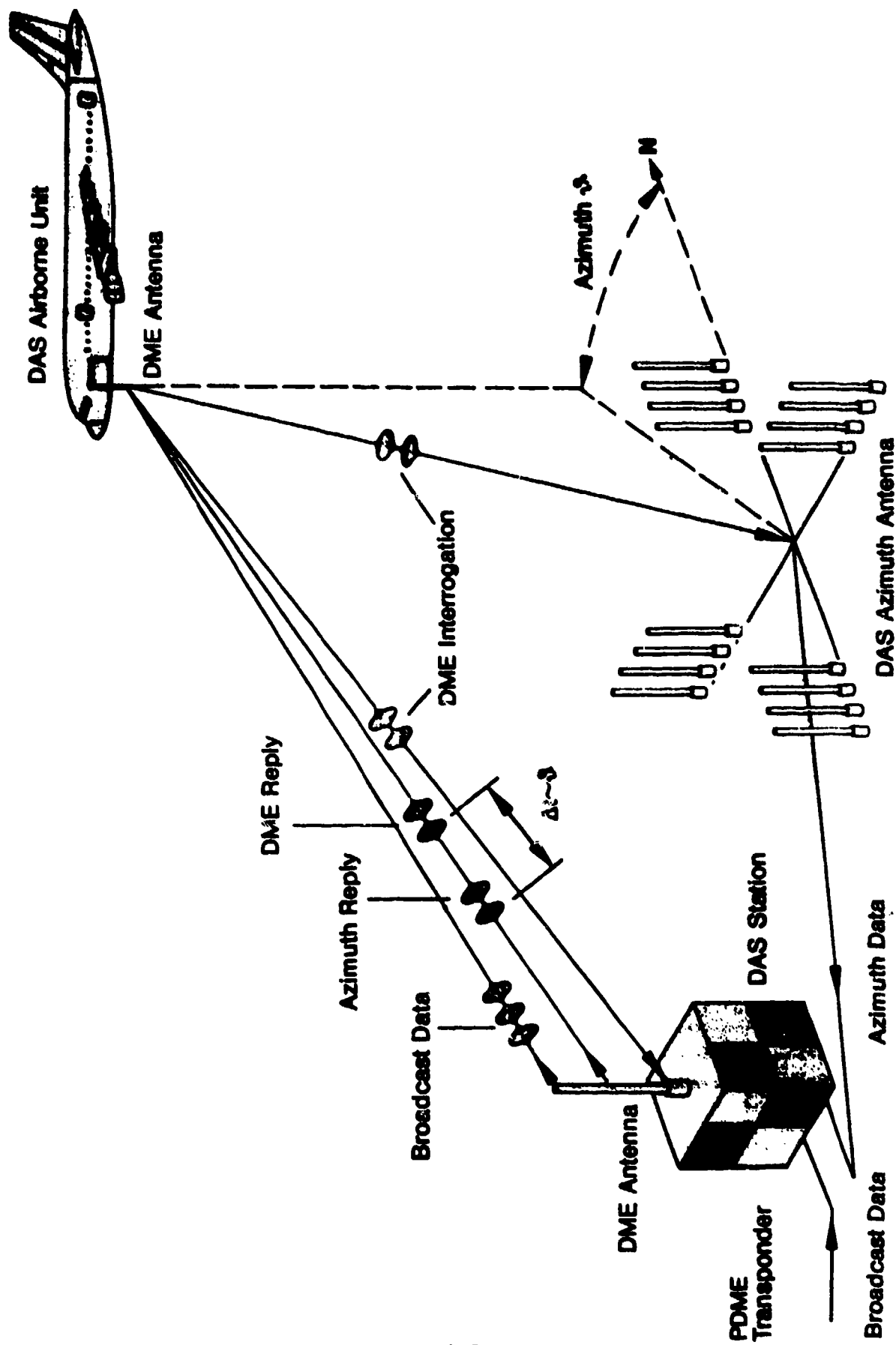
Fig. 4-2 DME Based Azimuth Landing System (DAS)

4-5

WAVE FRONT

$$\Delta\Phi = 2\pi\,\frac{d}{\lambda}\,\cos\gamma\,\sin\vartheta$$

INCOMING SIGNAL

STEP 1

RAW ANGLE ESTIMATE
AMBIGUITY RESOLUTION
WITHIN THE 6 ELEMENT
SUBGROUP

STEP 2

PHASE-AMBIGUITY RESOLUTION
OVER THE TOTAL LENGTH
OF THE ARRAY

$$\Delta\Phi = (\Phi_n - \Phi_1) + N \cdot 360°$$

STEP 3

FINAL DETERMINATION OF
PHASE DIFFERENCE $\Delta\Phi_{VL}$
FOR THE LINEAR ARRAY

STEP 4

COMBINATION OF THE RESULTS
OF BOTH LINEAR ARRAYS

$$\vartheta = \arctan\frac{\Delta\Phi_V}{\Delta\Phi_H}$$

$$\gamma = \arccos\frac{1}{2\pi}\,\frac{\lambda}{d}\sqrt{\Delta\Phi_V^2 + \Delta\Phi_H^2}$$
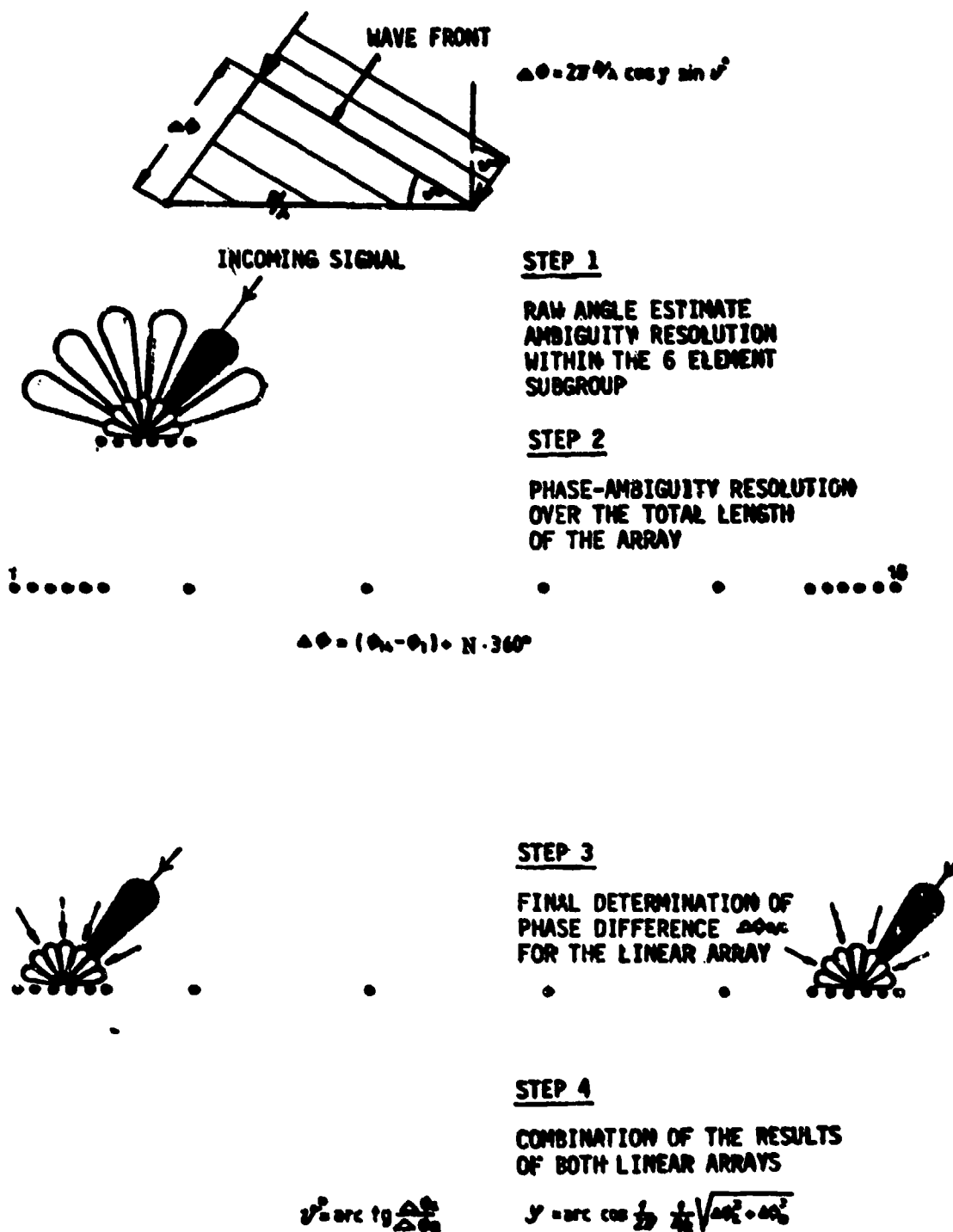
Fig. 4-3. DAS angle-measurement principles.

mitigate the effects of multipath with a separation angle greater than 10°
(rather than provide an improvement in system accuracy against inbeam
multipath).

In cases where the multipath angle is always on one side of the direct
signal angle, (e.g., in most systems that measure elevation angle, the multi-
path angle of arrival is typically less than that of the desired signal), the
"synthetic element" pattern can be offset to favor the direct signal. Methods
of this sort were investigated extensively in the context of the MLS assess-
ment and hence were not considered in the current study.

2. Signal Parameter Estimation

At various points in the DLS/DAS development process, suggestions were
made that one could utilize the amplitude and phase information received at N
antenna in a linear array to determine the parameters (complex amplitude and
angle of arrival) of $2N/3$ arriving plane waves. One such scheme applicable to
the case of three antennas and two plane waves was discussed by Bohm [10].

The general case of N antennas has been examined in some detail by Kupiec
[38]. He suggests that the solution for the case where $M=2N/3$ plane wave
signals are present can be efficiently obtained from the set of equations:

$$s_n = \sum_{m=1}^{M} p_m z_m^n ; \quad n=1, \ldots, N \tag{4-1}$$

$$\sum_{m=0}^{M} s_{n+m} a_m = 0 ; \quad n=1, \ldots, M/2 = N/3 \tag{4-2a}$$

$$\sum_{m=0}^{M} s_{n-m}^* a_m = 0; \quad n = M+1, \ldots, N \tag{4-2b}$$

where

p_m is the complex amplitude of a plane wave at angle θ_m

$z_m = \exp(j2\pi\xi \sin \theta_m)$

ξ = element spacing in wavelengths

s_n = signal at n^{th} element

$a_0 = 1$

and the remaining $\{a_i\}$ determine the coefficients of an Mth order polynomial $A(z)$ whose roots are the $\{z_m\}$. One first solves (4-2) for the unknown polynomial coefficients and computes the corresponding roots. The roots can then be substituted into Eq. (4-1) to determine the $\{p_m\}$. The system of equations in (4-2) may be interpreted as a forward/backward extension of Prony's method (see Chapter II).

Two problems exist with the algorithm described above:

(1) The existence of additive noise is ignored, and

(2) No procedure is specified for determining the proper number of plane waves directly from the array data.

The effects of noise on the above algorithm were simulated by Kupiec [38] for the case of two plane waves and a three-element array (this is the case considered by Bohm [10]). It was found that that the algorithm worked reasonably well at most separation angles (rms error < 0.15 beamwidth) except when the plane waves are nearly in (or out of) phase at the middle element of the array. When the inphase condition occurs, the system of equations represented by (4-2) becomes singular with the result that small amounts of noise can cause large errors (e.g., > 0.5 beamwidth).

Kupiec [38] suggests some methods for distinguishing when the received data lie near these singular points so that an alternative algorithm might be used. However, the results in Chapter III of this report suggest that the problem encountered by Kupiec is fundamental (i.e., no algorithm will work in an inphase condition if $M > N/2$).

Including noise in the problem formulation at the outset leads one to the notion that rather than assuming the received signals are exactly given by (4-1), one should attempt to seek a best approximation $\hat{s}$ to the observed data $\underline{r}$ by functions of the form of (4-1). Choosing a least-square approximation that minimizes

$$C = \sum_{n=1}^{N} |r_n - \hat{s}_n|^2 \tag{4-3}$$

leads to the maximum likelihood estimate discussed in Chapter III. It is worth noting that Kupiec [38] concluded that the use of (4-3) gave significantly better performance than (4-2) in the cases he simulated.

B.   Maximum Likelihood Parameter Estimate

In the preceding chapter, we found that the maximum likelihood (ML) parameter estimate involved maximizing a rather complicated function of the L signal directions. This result was practically important since a direct ML approach using Eq. (3-1) involves maximizing a function in 3L variables as opposed to L variables. Although theoretically interesting, we have not assessed this approach numerically in the present study due to uncertainty in

(1)   the extent to which the gradient search would converge to the ML estimates when a large model order is required, and

(2)   difficulties in determining what model order would be appropriate for a given situation if no side information was available.

Appendix E describes a preliminary assessment of the model order issue conducted by S. Dolinar in the early portion of the present study. Numerical evaluations of the ML estimate for the case of two or more plane-wave signals have been published by Birgenheier [36], Kupiec [37], Howard et al. [39], Nickel [35], Cantrell et al. [56], Trunk et al. [59], and Howard [57]. Table 4-1 summarizes the results of these various numerical investigations. Birgenheier also considers the problem of determining the proper model order given only the observation data.

TABLE 4-1

NUMERICAL DETERMINATION OF ML ANGLE OF ARRIVAL ESTIMATES

Researcher	# Plane Waves	# Array Elements	Comments
Kupiec [38]	2 3	3 6	Considered exact Prony as well as ME.
	11	6	10 small signals near horizon to simulate diffuse multipath.
Kupiec [37]	2,3	5	L band field measurements of reflection from flat terrain
	2	3	Three element results were significantly poorer than five element results. Three elements config. yielding errors $\leq$ .05 BW down to 0.4 BW separation angle.
Birgenheier [36]*	2,3	8	Shows rate of convergence for a number of cases.
	4	8	Converged for some, but no all signals.
	5	8	Did not converge.
Howard [57]	4	2	
Nickel [38]	2 3	3,5,9 ?	Shows log likelihood function contours and detection probability curves.
Cantrell, et al. [56]	2	3,21	ME estimates compared with CR bounds. In some cases, ME runs error was better than unbiased CR bound.
Trunk, et al. [59]	2	20	Shows prob. density function for cases where it is known that the two signals are symmetrical about array boresight. Bimodel densities arise at small separation angles (see Fig. 4-4).

*Considers order selection problem.

All above consider case of a single observation with a uniform spaced line array.

The numerical results of Trunk, Cantrell, and Gordon [59] provide inter-
esting insight into the problems likely to be encountered in estimating the
arrival angle of individual plane waves. They consider the special case of
two signals which are known to be symmetrical about the array boresight, but
with unknown amplitudes and relative phases. In this special case, one is
interested primarily in estimating the relative angle-of-arrival of the two
plane waves.

Figure 4-4 shows computed probability density functions (pdfs) of the ML
angle estimate for several angular wave separations and relative phases. In
all cases, the relative amplitude of the reflected (i.e., lower) plane wave
was -1 dB. At high SNR, the pdf of the ML estimate is roughly Gaussian[*] and
centered about the true angle, corresponding to a region where Cramer-Rao
bounds give a good estimate of estimation performance. At lower SNR, the pdf
develops an additional impulse function near the horizon, corresponding to
failure to resolve the two plane waves.

The impulse function which arises in Fig. 4-4 is of particular interest
since our experience to date with most practical high-resolution schemes (see
Chapter V) is that performance degradation at low SNR is primarily caused by
an inability to resolve the plane waves. This phenomenon is manifested as a
single angle estimate near the centroid of the two angles of arrival.

C.   Maximum Likelihood Method (MLM)

The MLM estimation technique had its genesis in seismic array beamforming
under conditions of directional interference [12]. The underlying problem is
to determine the minimum variance unbiased estimate of the complex signal from
a given angle θ subject to (complex) interference with a known covariance
matrix $\underline{Q}$. The signal is estimated by weighting the received signal at each
element and summing (see Fig. 4-5) such that the voltage gain in direction θ

---

[*]Although the pdf does not appear to be Gaussian, especially at lower SNR, it
is shown in [59] that the corresponding distribution function matches that of
a Gaussian function fairly well.
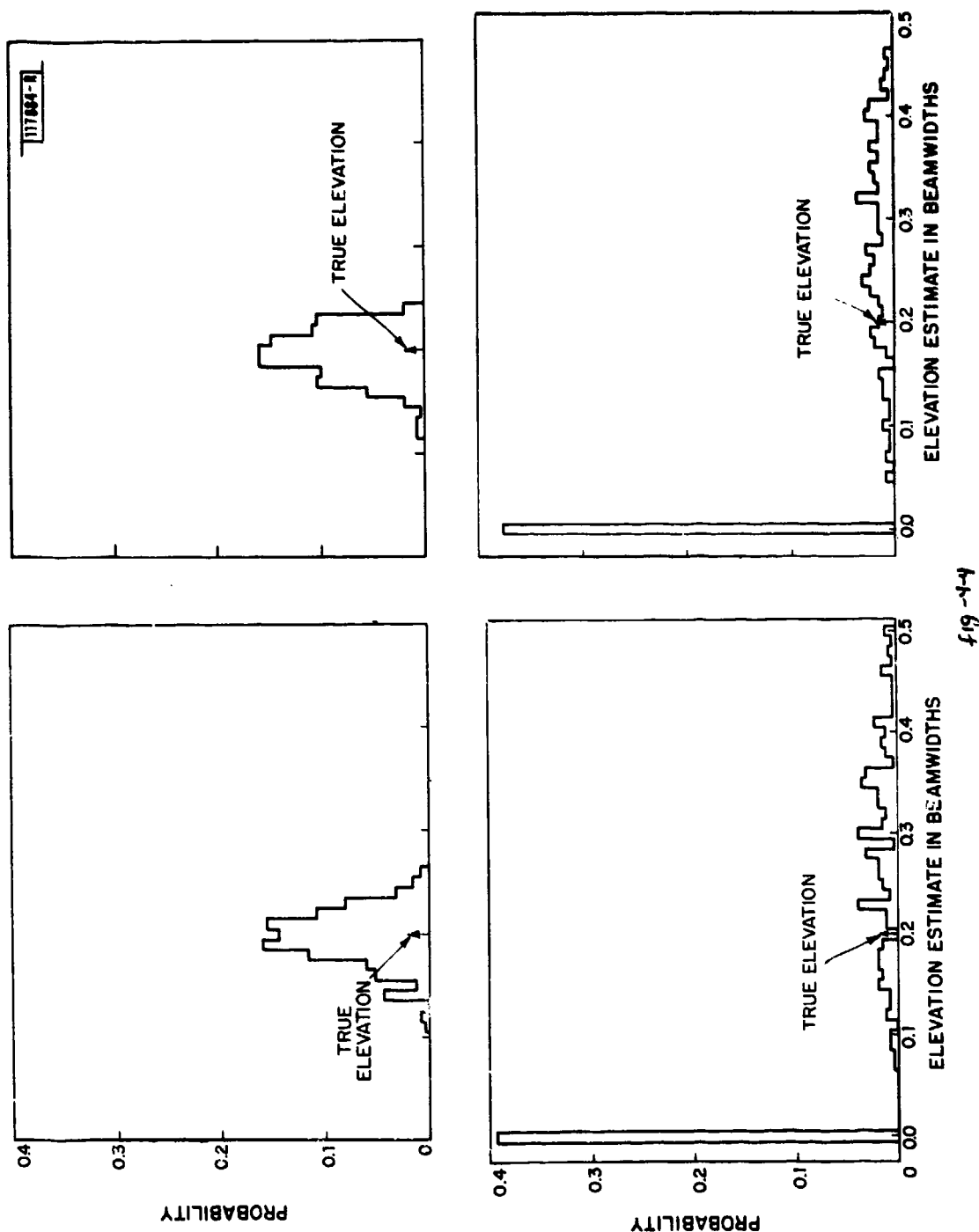
Fig. 4-4. Histograms of Maximum Likelihood Function for two symmetrical plane waves in additive noise (from Trunk et al. (59)).

4-12
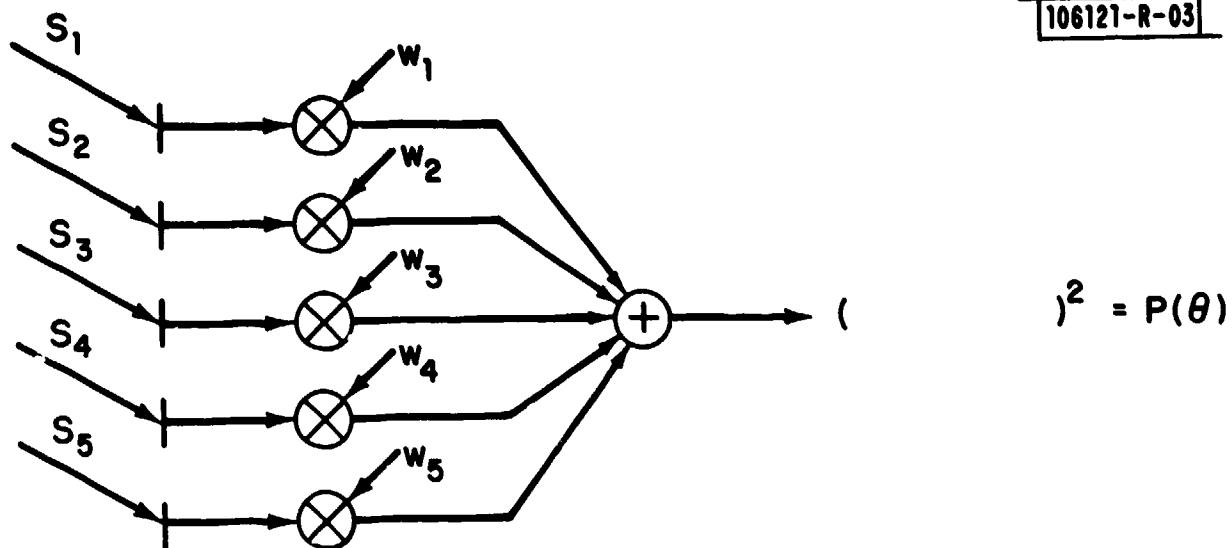
$$v_i = e^{jkx_i \sin\theta}$$

$$\text{FOR ML:} \quad \underline{w} = \frac{\underline{\underline{Q}}^{-1} \underline{v}}{\underline{v}^H \underline{\underline{Q}}^{-1} \underline{v}}$$

Q = ESTIMATED INTERFERENCE MATRIX

Fig. 4-5. Maximum likelihood spectrum estimation.

is held fixed at unity and the weight factors yield the minimum expected output power.

When the interference is Gaussian, the maximum likelihood estimate of the average power in a plane wave from angle θ is given by

$$P_{ML}(\theta) = <|\underline{w}^H\underline{r}|^2> = \underline{w}^H\hat{\underline{R}}\,\underline{w}$$

$$= \underline{v}^H\underline{Q}^{-1}\hat{\underline{R}}\,\underline{Q}^{-1}\underline{v}/(\underline{v}^H\underline{Q}^{-1}\underline{v}) \qquad (4\text{-}4)$$

where $\underline{v} = \underline{v}(\theta)$ is the received signal vector corresponding to a unit plane wave from angle θ, $\hat{\underline{R}}$ is the sample covariance matrix, and $\underline{w} = \underline{w}(\theta)$ is the element weighting vector given by

$$\underline{w} = \underline{Q}^{-1}\,\underline{v}/(\underline{v}^H\underline{Q}^{-1}\underline{v})$$

If the desired signal is absent, the expected value of the sample covariance matrix is $\underline{Q}$, and the expected value of the ML power estimate becomes

$$E\{\hat{P}_{ML}(\theta)\} = [\underline{v}^H(\theta)\underline{Q}^{-1}\underline{v}(\theta)]^{-1} \qquad . \qquad (4\text{-}5)$$

The right-hand side of (4-5) is also the variance of the ML estimate of the complex signal.

When the interfering signals are multipath, $\underline{Q}$ cannot be measured independently of the desired signal. This suggests substituting the sample covariance matrix $\hat{\underline{R}}$ for $\underline{Q}$. The resulting MLM angular power spectrum estimate is

$$P_{MLM}(\theta) = (\underline{v}^H\,\hat{\underline{R}}^{-1}\underline{v})^{-1} \qquad . \qquad (4\text{-}6)$$

The resolution performance of the MLM technique was first considered by Cox [3] for the case where the receiver has an accurate estimate of the ensemble covariance for uncorrelated signal processes. Cox shows that in such a case, the MLM technique can resolve signals (in the sense of yielding separate spectral peaks) which are less than a beamwidth apart provided that adequate

signal to noise ratio (SNR) is available. Figure 4-6 shows the minimum required SNR for two equal amplitude signals as a function of the angular separation in beamwidths. An interesting feature of these results is that resolution of signals with very different amplitudes (e.g., 30 dB) requires only slightly greater power as indicated by the second curve in Fig. 4-6. The depth of the dip between the peaks at which one declares resolution to be achieved plays a secondary role in generating the curve of Fig. 4-6 as is shown in Fig. 4-7.

The low sidelobes of the MLM spectrum in the incoherent case are achieved by virtue of the element weighting vector $\underline{w}(\theta)$ which places nulls at the angles of all emitters. When θ coincides with a plane wave arrival direction, the unity gain constraint prevents a null from being formed at θ. The resolution of the MLM spectrum is ultimately limited by the fact that two nulls can be placed very close together (on either side of θ) only at the expense of a beam pattern with a very large gain at other angles (see Fig. 4-8), a situation which requires large element weights. Since the output power due to receiver noise is proportional to $|\underline{w}(\theta)|^2$, a finite SNR will ultimately prevent nulls from being formed at the "nearby" plane wave directions.

In the case of coherent signals (e.g., multipath), the problems with MLM spectrum estimates become more severe since the array output power is no longer minimized by steering nulls to the signal angles other than θ. Rather, it turns out (see App. F) that minimum array output power is achieved by choosing a weighting vector $\underline{w}(\theta)$ that causes the desired and undesired signals to be cancelled at the array output. The resulting MLM spectrum (see Fig. 4-9) is similar to the classical Fourier (beamsum) spectrum except that it has a much higher noise level. When only a small number of observations are available, statistical fluctuations in the background noise level can lead to erroneous MLM estimates at a SNR which is much higher than those which would yield problems with classical spectrum analysis methods.

When correlated signals (e.g., multipath) are present, the signal cross terms contribute to the sample covariance $\hat{\underline{R}}$ in a way that depends on the relative rf phases between the signal paths. In contrast, the signal cross
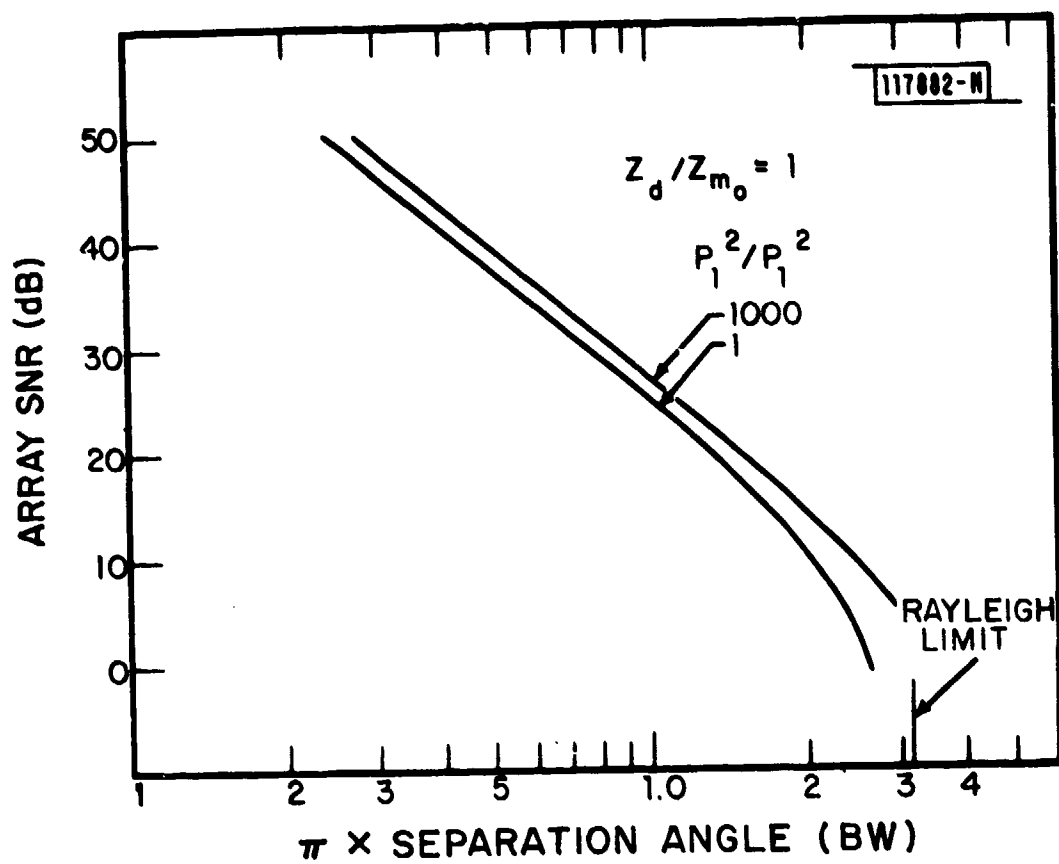
Fig. 4-6. Comparison of sufficient condition for resolution of equal and unequal sources with a line array (from Cox (3)).

4-16
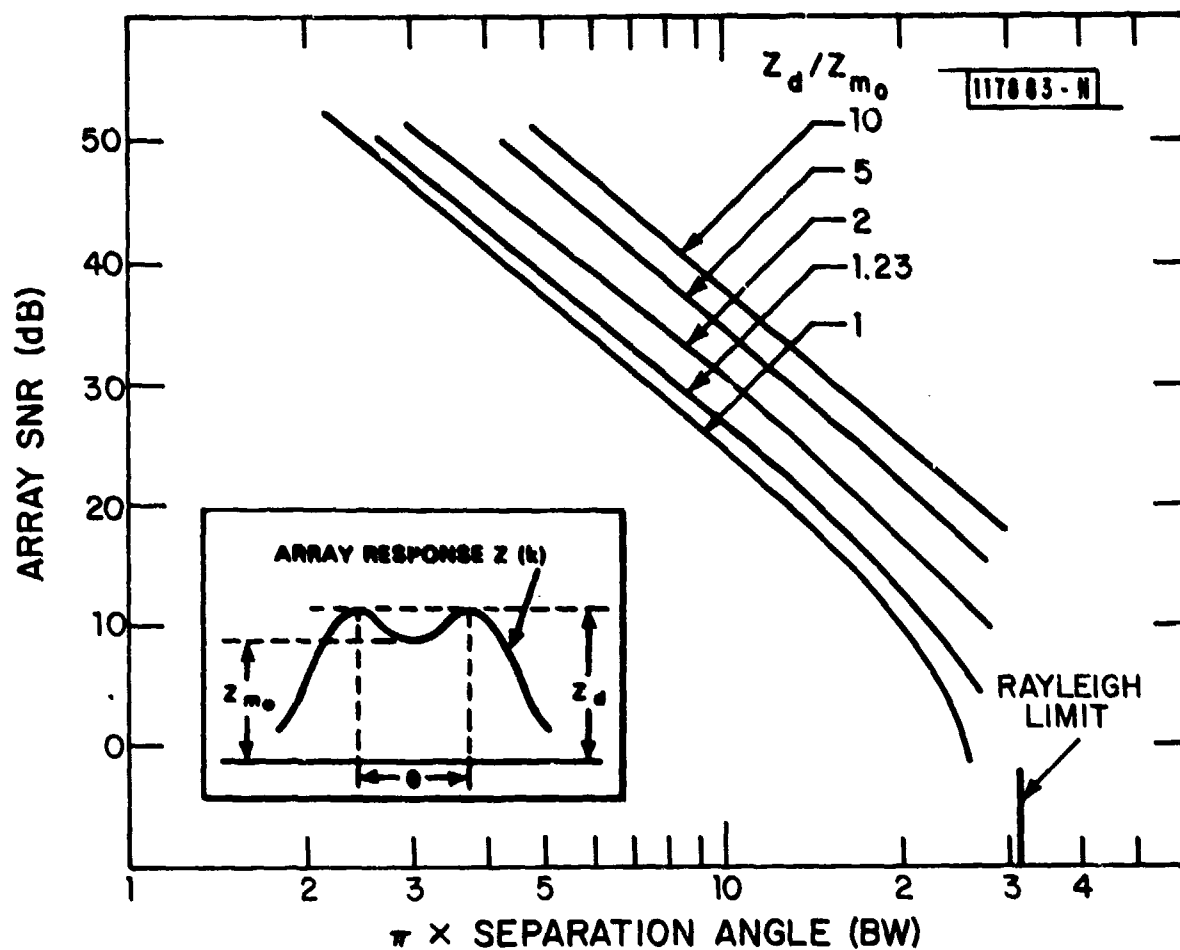
Fig. 4-7. Required value of the array output signal-to-noise ratio versus normalized angular separation for resolution by a line array for various levels of the "on target" to "mid point" response ratio (from Cox (3)).
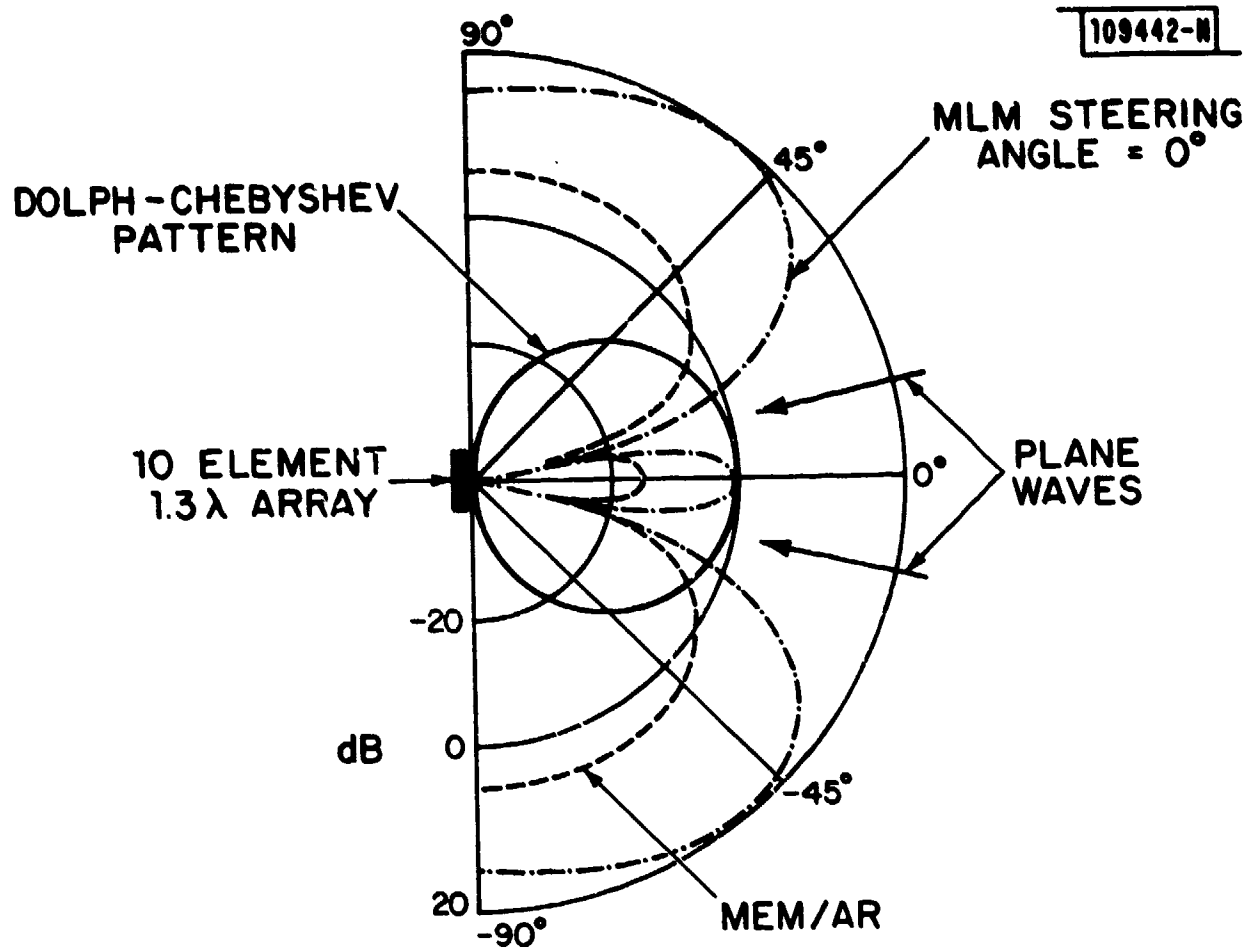
Fig. 4-8. MLM array gain at angle between arrival angles of two plane waves.

4-18
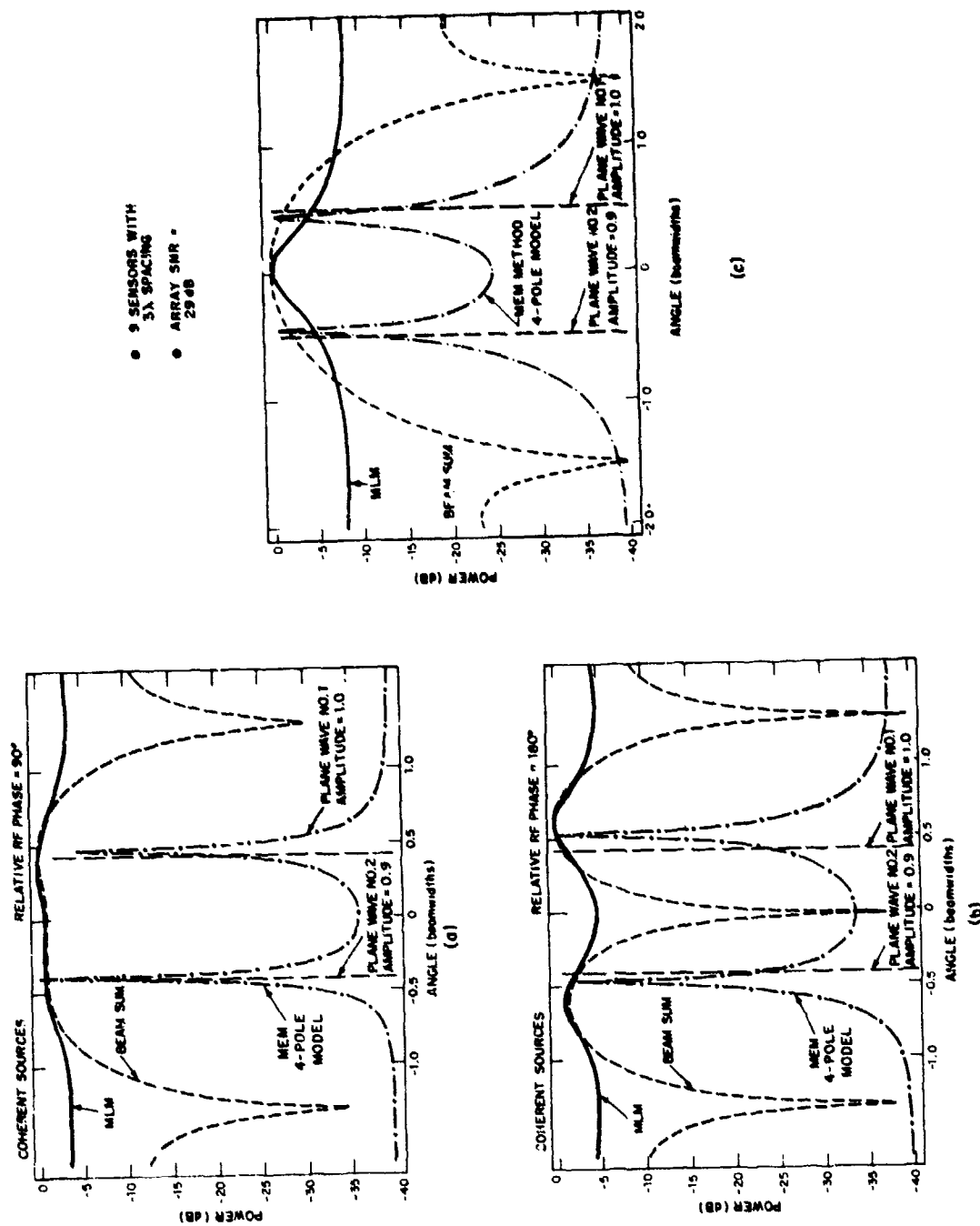
Fig. 4-9. MIM spectra without spatial averaging.

4-19

terms are zero in the uncorrelated signal ensemble covariance. One way of reducing the effect of the signal cross terms is to spatially average the raw covariance samples to yield a smoothed covariance estimate. Averaging the entries of $\hat{R}$ that correspond to any given separation between array elements yields a Toeplitz covariance estimate (i.e., one in which all terms along a diagonal are equal). Although such estimates have been widely used in maximum entropy/autoregressive modeling (see section D of this chapter), they have been found to behave very erratically in coherent signal environments.

The success of the "modified covariance" estimate in the case of the extended Prony/maximum entropy techniques suggested that it might also improve the performance of the MLM technique. In particular we considered the estimate

$$\hat{P}_{MLM}(\theta) = (\underline{v}^H \hat{\underline{R}}_M^{-1} \underline{v})^{-1} \tag{4-7}$$

where $\hat{\underline{R}}_M$ is the "modified covariance" estimate of the uncorrelated signal ensemble covariance. Equation (4-7) can be interpreted as follows:

1. The estimate of the ensemble covariance for an N-M+1 element array, $\underline{R}_M$ is obtained by averaging the raw covariance estimates from M overlapping subarrays.

2. The estimated power at angle θ for the ith subarray is obtained by weighting the raw received data by

$$\underline{w}^H(\theta) = \hat{\underline{v}}^H \hat{\underline{R}}^{-1} / (\underline{v}_{-M}^H \hat{R}^{-1} \underline{v}_{-M}) \tag{4-8}$$

so as to give

$$\hat{P}_i(\theta) = (\underline{v}^H \hat{\underline{R}}_M^{-1} \hat{\underline{R}}_i \hat{\underline{R}}_M^{-1} \underline{v}) / (\underline{v}^H \hat{\underline{R}}_M^{-1} \underline{v})^2 \quad . \tag{4-9}$$

3. The final estimate of power at angle θ is obtained by averaging the subarray power estimates, i.e.,

$$\hat{P}_{MLM}(\theta) = \frac{1}{M} \sum_{i=1}^{M} \hat{P}_i(\theta) \quad . \tag{4-10}$$

D.   Maximum Entropy/Autoregressive Spectral Estimation

The use of maximum entropy (ME)/autoregressive (AR) methods for achieving higher resolution spectral estimates can be motivated by a variety of argu-

ments (for example, see Appendix D). The most direct approach is to assume that the observed signals have been generated by the equation

$$r_n = \sum_{m=1}^{M} a_m r_{n-m} + \epsilon_n \qquad (4\text{-}11)$$

where ϵ_n is a zero mean, uncorrelated sequence with variance σ_ϵ^2. The objective of the ME/AR techniques is to estimate the $\{a_m\}$ and the σ_ϵ^2 given the measured data. This is accomplished as follows:

(i) choose the weights $\{a_m\}$ so as to make the residual sequence

$$u_n = r_n - \hat{r}_n = r_n - \sum_{m=1}^{M} a_m r_{n-m} \qquad (4\text{-}12)$$

spectrally white and of minimum variance.

(ii) then

$$U(z) = \sum u_n z^{-n} = \sum |r_n - \sum a_m r_{n-m}| z^{-\ell}$$

$$= W(z)\, R(z)$$

where $W(z)$ represents the z-transform of the filter weights

$$\underline{w}^T = |1, -a_1, -a_2, \ldots -a_M| \qquad . \qquad (4\text{-}13)$$

(iii) if the sequence $\{u_n\}$ is indeed white, then

$$|U(z)|^2 = \rho_M$$

and we have the spectrum estimate

$$|R(z)|^2 = \frac{\rho_M}{|W(z)|^2} \qquad . \qquad (4\text{-}14)$$

Another interpretation of ME/AR estimation is in terms of adaptive null-
ing beam patterns [1]. The ME/AR minimization procedure applied to the case
M=L can be viewed as determining an array pattern* which minimizes the array
output subject to the constraint that one of coefficients must be unity. This
is achieved by having nulls at the angles corresponding to plane wave arri-
vals, as illustrated in Fig. 4-10. Also shown in Fig. 4-10 is the
corresponding -30 dB sidelobe Dolph-Chebyshev array pattern as well as the MLM
pattern for estimating the power at an angle midway between the plane wave
arrival angles. The patterns in Fig. 4-19 have not been normalized so that
one can see that although the MEM and MLM patterns both have nulls at the
plane wave arrival angles, the MLM pattern gain is much higher at wide angles
due to the unity gain constraint at the steering angle.

The various ME/AR techniques differ principally in the method by which
the weights are to be determined. If the various plane-wave signals are
uncorrelated and the ensemble covariance $\underline{R}$ is given, it can be shown (see
Appendix D) that the weights $\underline{w}$ and the prediction error ρ_M can be determined
from the Weiner equation

$$\underline{R}\,\underline{w} = [\rho_M \; 0 \; 0 \; \ldots \; 0]^T \quad . \tag{4-15}$$

where $\underline{R}$ contains the first L lags of the correlation function along the main
diagonals. In practice, the available information usually consists of K data
records, $\underline{r}(k)$, k=1, . . . K. When K > 1, a standard practice is to construct
the sample covariance matrix $\hat{\underline{R}}(K)$ as discussed in Chapter II. When the number
of records K is less than the record length L, (i.e., each record $\underline{r}(k)$ is
assumed to be an L-dimensional vector) the (LxL) covariance estimate $\hat{\underline{R}}(K)$ is
singular. In the singular case, either or both of the spatial averaging
operations discussed in Chapter II may be employed with good prospects for

_____ __

*If the signals are uncorrelated and a large number of time observations are
available, one can set M = L = 1 in which case the nulling array coincides
with the physical array. If coherent signals are encountered, then the
nulling array corresponds to the "averaged" subaperture discussed previously.

DOLPH-CHEBYSHEV
PATTERN
(-30 dB Sidelobes)

45°

90°

10 ELEMENT
1.8λ ARRAY

PLANE WAVE
ARRIVAL
ANGLES

0°

-20

dB   0

MLM PATTERN
FOR 0° STEERING
VECTOR

-45°

20
-90°
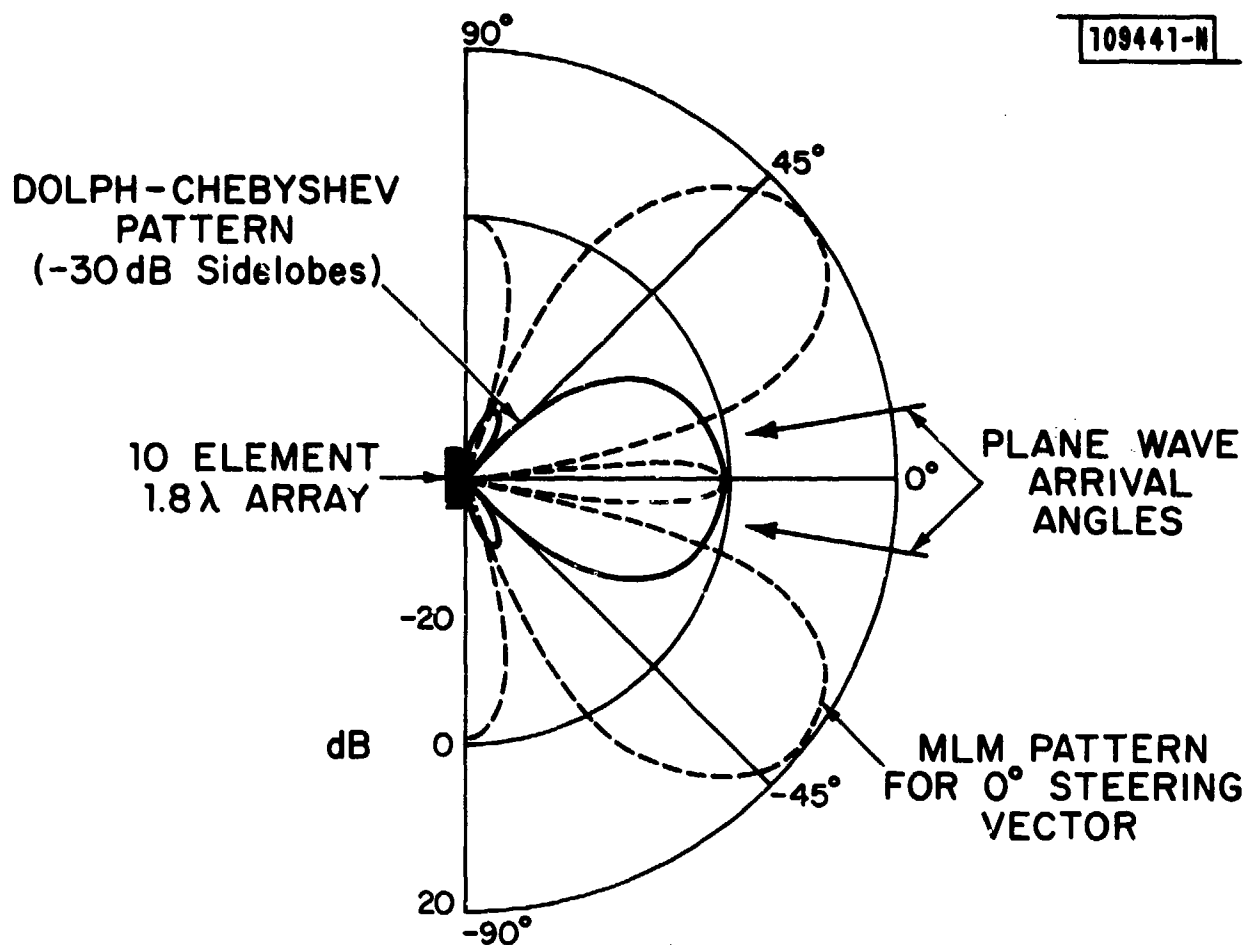
Fig. 4-10. Comparison of array patterns.

arriving at a non-singular covariance estimate. However, the spatial averaging operations discussed there do not usually yield a Toeplitz matrix.

The Toeplitz requirement may be imposed for one of two reasons. The first argument centers on the computational costs incurred in solving the Wiener equation (4-15) with $\hat{R}$ used in place of R for the (estimated) error filter weights $\underline{w}$. Usually, the first element of $\underline{w}$ is constrained to be unity, and the prediction error (variance) ρ_M constitutes the remaining unknown. The so-called "modern" Levinson recursion [62] solves (4-11) extremely efficiently, provided $\hat{R}$ is a Toeplitz (i.e., a correlation) matrix. The second argument hinges on the minimum-phase-property of the filter derived from the weights. In particular, if $\hat{R}$ is a positive-definite correlation matrix and $\underline{w}$ is the (unique) solution of (4-15), then the poles and zeroes of the z-transform of the elements of $\underline{w}$ all lie inside the unit circle. A filter with this property is said to be minimum-phase [60]. Obviously, a minimum-phase filter has a minimum-phase (and hence stable) inverse.

In a few specialized applications, the stability of the inverse filter is crucial (e.g., speech synthesis). However, many potential applications do not involve sample function generation, and the stability issue becomes relatively unimportant. Of course, computational costs are nearly always germane and due consideration should be given to especially efficient procedures. Thus it is unfortunate, but hardly surprising, that the standard procedures for obtaining correlation estimates do not work particularly well in a coherent signal environment. In fact, our simulation experience indicates that correlation methods generally perform worse than covariance methods. We say this without prejudice towards the "modified correlation" method described below (and as yet untested).

Correlation Estimates

A popular method for generating correlation (lag) estimates from an LxL sample covariance matrix $\hat{R}$ is to first sum the entries on each diagonal and then divide the sums by L, i.e.,

$$\hat{c}_\ell \overset{\Delta}{=} \frac{1}{L} \sum_{m=1}^{L-\ell} \hat{R}_{m+\ell,m} \quad ; \quad \ell = 0, 1, \ldots, L-1 \qquad (4\text{-}16)$$

It is a well-known fact that these estimates are always biased. Even when $E\{\hat{R}_{mn}\} = c_{m-n}$, we find that the expected value of $\hat{c}_\ell$ is

$$E\{\hat{c}_\ell\} = (1 - \ell/L)c_\ell \quad ; \quad \ell \geq 0$$

Evidently, the bias can be removed by a simple modification, i.e., the un-biased estimates

$$\hat{c}_\ell' \overset{\Delta}{=} \frac{1}{L-\ell} \sum_{m=1}^{L-\ell} \hat{R}_{m+\ell,m} \qquad (4\text{-}17)$$

are generated by computing the average value along each diagonal. Unfortunately, the (Hermitian) Toeplitz matrix constructed from the lag estimates in (4-17) may not be positive semi-definite. Since (4-16) is biased, neither of these estimates is recommended.

An alternative method for generating correlation estimates was suggested recently by Ulrych and Ooe [40]. In terms of a spatially-averaged sample co-variance matrix of the general form $_2\hat{R}_M(K)$ with $M < L$ (see Chapter II), spa-tially averaged lag estimates $\hat{\underline{c}} \overset{\Delta}{=} (\hat{c}_0, \hat{c}_1, \ldots, \hat{c}_{L-M})^T$ are extracted from the first <u>column</u> of $_2\hat{R}_M(K)$. Since spatial averaging only "shrinks" Toeplitz matrices, it follows that the spatially averaged lag estimates are <u>unbiased</u>. Ulrych and Ooe also claim that the lag matrix constructed from these estimates is non-negative definite.

The Burg Algorithm [61]

The popular Burg algorithm does not require lag estimates. In fact, it is usually said that Burg derives the error filter weights directly from the "data". As we shall see, the "data" generally takes the form of a covariance estimate.

The Burg error filter weights are computed from an iterative relationship of the form

$$\underline{w}_{n+1} = \underline{w}_{n+} - \hat{k}_n \, (\underline{w}_{n+})^\dagger \qquad\qquad (4\text{-}18)$$

The ("Levinson") recursion (4-18) starts with

$$\underline{w}_{0+} \stackrel{\Delta}{=} \begin{bmatrix} 1 \\ 0 \end{bmatrix}$$

and, in general,

$$\underline{x}_+ \stackrel{\Delta}{=} \begin{bmatrix} \underline{x} \\ 0 \end{bmatrix}$$

is the "zero-extension" of an arbitrary vector $\underline{x}$. The superscript "$\dagger$" in (4-18) is the exchange operator (see Chapter II), and the nth reflection coefficient $\hat{k}_n$ is estimated by minimizing the empirical n^{th} order prediction error. Since Burg treats the "original" data and the "exchanged" (i.e., reversed and conjugated) data on an equal basis, the nth reflection coefficient $\hat{k}_n$ may be obtained by solving*

$$\hat{k}_n \Leftarrow \min_{k_n} \, \underline{w}_{n+1}^H \; {}_2\hat{R}_{L-n} \; \underline{w}_{n+1} \qquad\qquad (4\text{-}19)$$

where ${}_2\hat{R}_{L-n}$ is the appropriate spatially-averaged (sample) covariance matrix, and $\underline{w}_{n+1}$ is given by (4-18) with $\hat{k}_n$ replaced by k_n.

The key to the Burg algorithm is that forward/backward averaging is always employed. Under this assumption, the reflection coefficients satisfy

$$|\hat{k}_n| < 1 \qquad\qquad (4\text{-}20)$$

provided the covariance estimates ${}_2\hat{R}_{L-n}$ are positive-definite at each stage of the recursion. Consequently, Burg filters always have the minimum-phase property.

---

*The procedure in (4-19) is applicable for any sample covariance matrix $R(K)$ or, as suggested by Burg [62], one may also consider generalized covariance estimates where some of the data (records) are given more weight than others.

Without going through the details, the solution to (4-19) may be written as

$$\hat{k}_n = \frac{2\left(\underline{w}_{n+}^{\dagger}, \underline{w}_{n+}\right)_{n+1}}{2\|\underline{w}_{n+}\|^2_{n+1}} \qquad (4-21)$$

where

$$_2\left(\underline{x}, \underline{y}\right)_{n+1} \overset{\Delta}{=} \underline{x}^H \,_2\hat{R}_{L-n}\, \underline{y} \qquad (4-22)$$

is the inner product induced by the covariance estimate $_2\hat{R}_{L-n}$, and

$$_2\|\underline{x}\|_{n+1} \overset{\Delta}{=} (\underline{x}, \underline{x})^{\frac{1}{2}}_{n+1} \qquad (4-23)$$

is the (usual) norm associated with the inner product in (4-23). Forward/backward averaging guarantees

$$_2\|\underline{x}\|_{n+1} = \,_2\|\underline{x}^{\dagger}\|_{n+1} \qquad (4-24)$$

for any (n+1)-dimensional vector $\underline{x}$. Consequently, an easy application of Schwartz's inequality establishes (4-20). The details of the argument may be found in Appendix D.

If only one data record is processed, the standard form of the Burg algorithm [63] calculates the reflection coefficient much more efficiently than (4-21). However, the standard "one-look" algorithm cannot be used to simultaneously "batch" process more than one record. In direction-finding applications, the number of observations (looks) may greatly exceed the length of an individual record. In this case, the "multiple-look" version of Burg's algorithm becomes computationally attractive. However, when the minimum-phase property is unimportant, the multiple-look Burg algorithm has no particular advantage over the "modified covariance" method.

Modified Covariance Method

The modified covariance method minimizes the quadratic form in (4-19) subject only to the constraint that the leading (first) element of $\underline{w}$ be unity. The only difference between this approach and the covariance method discussed in Chapter II is that the spatially-smoothed sample covariance matrix is also forward/backward averaged. Ulrych and Clayton [65] refer to this technique as a least squares method. The method has also been credited to Nuttal [66].

Loosely speaking, any of the methods discussed in this section for "solving" (4-15) are called "maximum entropy" methods. These techniques could just as well be called "autoregressive" methods or "linear prediction" methods. Given a "solution" $\underline{w}$ derived via any of these methods, two standard procedures for extracting the desired angle (frequency) estimates are often employed.

Angle Estimates

In the usual power spectral density approach, the discrete Fourier transform of the error filter weights (i.e., the elements of $\underline{w}$) is computed on a very fine grid. In adaptive array terminology, the magnitude of this function is referred to as the "adapted pattern". Local minima or "nulls" in the adapted pattern correspond to potential angle estimates. The "maximum entropy" estimate of the power spectral density is generated by squaring and inverting the adapted pattern, and the "nulls" in the latter correspond to the "peaks" of the power spectral density estimate as is illustrated in Fig. 4-9. This approach is referred to here as the maximum entropy method (MEM).

An alternative "line spectrum" approach calculates the roots of the polynomial with coefficients given by the error filter weights ($\underline{w}$). Taking the argument of these complex roots provides M potential angle (frequency) estimates. This technique is called the Autoregressive Root Method (ARM). Spurious ARM estimates are rejected by computing a power estimate for each potential angle estimate and comparing the power estimates to an appropriate (noise) threshold. An effective technique for computing the power estimates is discussed in the next chapter.

E. Auto Regressive-Moving Average (ARMA) Spectrum Modeling

It can be shown [87] that the sum of one or more plane wave signals in the presence of additive white noise can be modeled by an ARMA process of the form:

$$r_m - \sum_{i=1}^{L} a_i r_{m-i} = \epsilon_n - \sum_{i=1}^{M} b_i \epsilon_{n-i} \qquad (4-25)$$

where ϵ_t is a zero-mean uncorrelated sequence with variance σ_ϵ^2. Moreover, the model of (4-25) can be shown to also remain appropriate when diffuse multipath is present. Since pure autoregressive (AR) models [i.e., M=0 in (4-25)] only approximate the spectrum of these processes as the number of poles (L) approaches infinity, it seems reasonable that one could achieve better results by using an ARMA model. In this section, we outline two ARMA modeling approaches which were investigated in the course of the current study.

1. Unconditional Least Squares Estimation

The determination of least squares estimates for the ARMA parameters $\{a_i\}$ and $\{b_i\}$ when all variables are real has been described by Box and Jenkins [22]. However, the case of complex $\{a_i\}$, $\{b_i\}$, $\{r_n\}$ and $\{e_t\}$ had not been previously accomplished when our work was carried out. The basic idea is to choose the $\{a_i\}$ and $\{b_i\}$ to minimize the residuals

$$I = \sum_{t=1}^{N} |u_t|^2 \qquad (4-26)$$

where
$$u_t = x_t - \sum_{i=1}^{L} a_i x_{t-i} + \sum_{i=1}^{M} b_i u_{t-i} \qquad (4-27)$$

$\{r_n\}$ are the complex received data $n = 1$ to N.

This is accomplished in several steps:

(1) estimate an initial set of values for $\{a_i\}$ and $\{b_i\}$ by some appropriate algorithm. The particular algorithm adopted in our case was the modified Yule-Walker procedure due to Kaveh [21] which will be discussed subsequently.

(2) Minimize the objective function with respect to the ARMA
parameters using a gradient descent method. This is done
by computing the gradient of the objective function (4-
26) by finite differences, e.g.,

$$\frac{\partial I}{\partial a_1} \approx I(a_1 + \delta, a_2, \ldots) - I(a_1, a_2, \ldots b)/\delta$$

Several algorithms for minimizing I given the current $\underline{a}$
and $\underline{b}$ values and gradient were investigated. The most
stable results were found by using the Int. Math and Sci.
Library (IMSL) [24] conjugate gradient subroutine. This
routine operates on real variables only. Therefore, all
variables and data were broken down into real and
imaginary parts and the optimization then proceeded
forward as with real data.*

One of the significant problems in accomplishing step 2 arises from the
need to avoid end effects on the generation of the residuals and therefore on
the accuracy of the estimates. We note from (4-13) that the M previous resid-
uals and L previous values of x_t are used to generate u_t. If we consider only
t>L (as is the case with the ME/AR "modified covariance" algorithm), then
there are enough values of x_t to use in (4-13). However, the $\{u_i\}$ for i<t do
not represent measured quantities. Rather, one must estimate them from the
data. Unfortunately, when applying (4-13) repeatedly to generate the $\{u_{t-1}\}$,
we find that values of x_t for t<0 are required to generate the u_i when i < L.

The method used to alleviate (but not eliminate) this difficulty is to
use backforcasting of the data before the first observation by using (4-13)
and the backward residual equation

$$w_n = r_n - \sum_{i=1}^{L} a_i^* r_{n+i} - \sum_{i=1}^{M} b_i^* w_{n+i} \qquad (4-28)$$

Eq. (4.28) is used to generate w_n starting at n=N for

t=N, N-1, . . . 1   with $w_n = 0$ $n \leq 0$

---

*As an outgrowth of his work on the least squares algorithm described here,
Y.L. Chu is considering the use of a constrained search algorithm based on
pole-zero magnitudes and angles as a part of his S.B. thesis at M.I.T. [53].

Furthermore, we assume that

$$w_n = x_n = 0 \quad \text{for } t > N \tag{4-29}$$

The backforcasting procedure above was necessitated by the existance of the zeros in the assumed data transfer function whereas with the ME/AR modeling these problems did not arise. Figure 4-11 illustrates the differences between these two approaches in terms of determining a whitening filter for the received data.

Because of the backforcasting problem and the need for a computationally demanding numerical search procedure, the least squares algorithm described above is not viewed as currently practical for improved real time system accuracy. Rather, we view it as furnishing a benchmark which one can use to assess the performance loss with other more practical ARMA techniques and obtain insight into the potential available with an ARMA modeling approach vis a vis other approaches.

2. Modified Yule-Walker (MYW) ARMA Estimator

The difficult numerical problems associated with the general least squares estimation procedure motivated the investigation with suboptimal methods for estimating the ARMA parameters. There have been a number of suboptimal estimators described recently in the literature [25-26]. We chose to focus on the MYW estimator considered by Kaveh[*] [21] since that method had been shown to be superior to some of the other approaches and comparable to the Box-Jenkins in some limited cases [23].

The basic idea is that x_t is assumed to have an autocorrelation function r_t with power spectrum

$$S_x(z) = \frac{B(z) \; B^*|(z^*)^{-1}|}{D(z) \; D^*|(z^*)^{-1}|} \tag{4-30}$$

---

[*]Prof. M. Kaveh of the Univ. of Minnesota was a consultant to Lincoln during the course of these studies.
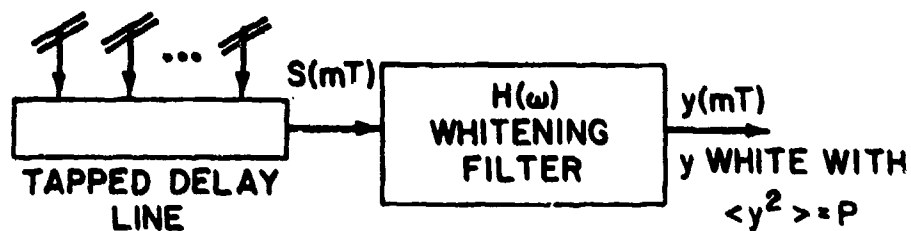
$$P(\theta) = P|H(\omega)|^{-2} \qquad \omega = k\delta \sin\theta/T$$

For ME/AR

$$H(z) = 1 - \sum_{k=1}^{M} a_k z^{-k} \text{ has only zeros}$$

h(kT) input transient is M samples long

For ARMA

$$H(z) = \Sigma b_k z^{-k} / \Sigma a_k z^{-k} \text{ has poles and zeros}$$

h(kT) input transient may last for all L samples

Fig. 4-11. Digital whitening filter interpretation of AR/ME and ARMA estimation.

where $d_0 = 1$ and $d_i = -a_i$ for $i \geq 1$. Since only quadratic functions of the $\{b_i\}$ are required, Kaveh suggests instead determining the $\{c_i\}$ such that

$$\sum_{k=-M}^{M} c_k z^k = B(z) B^* |(z^*)^{-1}| \ . \tag{4-31}$$

Eq. (4-31) can be interpreted as the spectrum of the residual process $\{u_t\}$ with the $\{\tilde{c}_k\}$ the estimated autocorrelation function of $\{u_t\}$.

The AR coefficients are determined from the extended Yule-Walker equation

$$\underline{C}_M \, \underline{a} = \underline{f}_M \tag{4-32}$$

where

$\underline{a}$ is the vector of the $\{a_i\}$

$$\underline{C}_M = \begin{bmatrix} c_M & c_{M-1} & \cdots & c_1 \\ c_{M+1} & c_M & \cdots & c_2 \\ & & & \\ c_{2M-1} & c_{2M-2} & \cdots & c_M \end{bmatrix} \tag{4-33}$$

$$\underline{f}_M = \begin{bmatrix} c_{M+1} \\ \cdot \\ \\ \cdot \\ c_{2M} \end{bmatrix} \tag{4-34}$$

We observe that when M=0 (corresponding to a pure ME/AR model), the equations in (4-31) to (4-34) reduce to the classical Yule-Walker equations.

Following the calculation of the $\underline{a}$ by inverting $\underline{C}_M$, Kaveh has shown [21] that the $\{\tilde{c}_k\}$ can be found as

$$\tilde{c}_k = \tilde{c}_{-k}^* = \sum_{i=0}^{M} \sum_{j=0}^{M} a_i \, a_j^* \, c_{|i-j-k|} \tag{4-35}$$

It should be noted that the algorithm above is non recursive in that the resulting $\tilde{c}$ (or, $\underline{b}$) estimates are not used to improve on the $\underline{a}$ estimates. This is because the effects of the zeros on the observed data were presumably taken into effect by using the higher order lags of c_k in estimating the $\underline{a}$.

A significant problem in using this technique is the method by which c_k is estimated. Kaveh's original paper [21] suggested the biased estimator of eq. (4-16). Problems were encountered initially with negative power estimates using this estimator, and some simulations were also carried out using the unbiased estimate without any significant improvement. More recently, Bruzzone and Kaveh [23] have suggested that the $\{c_k\}$ should be linearly tapered, e.g., by

$$T_k = 1 - k/K \qquad\qquad K > M \qquad\qquad (4-36)$$

to eliminate the negative power estimates at the expense of a (hopefully slight) loss in resolution.

3. Order Determination

One of the important issues in ARMA estimation is the choice of model order for the estimated spectrum. This is complicated by the fact that in general, one could choose L and M independently. However, the received signal in most ATC applications should be adequately modeled as a sum of AR processes and white noise where:

(a) the plane waves corresponds to a first order AR process with a pole on the unit circle, and

(b) diffuse signals correspond to a first order process with a pole inside the unit circle.

If the sum process spectrum has L poles, then it also will in general have M=L zeros. Consequently, our studies have focused on the case M=L.

Several quantitative criteria have been suggested for determining the optimum ARMA order for a given data set:

1. the Akaike generalized AIC test [17]

$$AIC(L) = N \ln \sigma^2 + 2\alpha \ (M+L) \tag{4-37}$$

where σ^2 is the Lth order residual power and
$1 < \alpha < 3$ ($\alpha = 1$ is normally used), and

2. the Akaike final prediction error FPE criteria which
    seeks to minimize a weighted residual sum of squares.

One problem with both of these criteria is that the whitening filter "memory"
corresponding to the $\{b_i\}$ parameters extends to times before the data has
arrived. Thus, the measured σ^2 for short data sets (i.e., a small number of
antenna elements) may heavily reflect the whitening filter transient response
as opposed to the steady state response.

It was not possible to carry out a detailed study of these order determi-
nation criteria within the time frame of the present study. Rather, our
studies considered simulation results with the correct choice for the given
simulation data and with values of L which were too large or too small so as
to gain some feeling for the model sensitivity to the choice of L.

F.   Eigenvector Methods

Much of the current interest in eigenvector methods can be traced to the
work of Pisarenko. In [67], he re-examined Caratheodory's trigonometric
moment problem, as extended by Szego [68]. This classic problem applies
directly to the incoherent signal case when direction-finding with an ideal
array. Reddi [69] considered a general signal-in-space covariance matrix (see
Chapter II), but his treatment of the coherent (singular) case was rather
incomplete. Schmidt [30] approached the direction-finding problem in a gener-
al manner and obtained the fundamental orthogonality result in Chapter II.
However, Schmidt did not discuss the important singular case that arises in
specular multipath environments.

Pisarenko's Method

The underlying trigonometric moment problem (theorem) permits the follow-
ing assertion: Given a starting sequence $\{c_0, c_1, \ldots, c_N\}$ of any (dis-
crete) correlation function, a consistent spectral representation can be

constructed from M "line" spectral components (i.e., frequency/power pairs) and a real constant (i.e., a "white" noise power level). Moreover, $M < N$ and the parameters of the representation, including M, are uniquely determined by the given correlation lags.

In order to calculate the frequencies and the power levels from the lags, Pisarenko arranges the latter in a correlation matrix $\underline{C}$. In terms of the "direction matrix" notation developed in Chapter II, the Pisarenko decomposition may be written as

$$\underline{C} = \underline{V}(\underline{\alpha}) \; \underline{\Psi} \; \underline{V}^H(\underline{\alpha}) + \nu \underline{I} \qquad (4\text{-}38)$$

where $\underline{\alpha}$ represents the (angular) frequencies, $\underline{\Psi}$ is a (real) positive-definite diagonal matrix containing the powers for the frequency components, and ν is the white noise level. It follows that ν is the minimum eigenvalue of $\underline{C}$. When ν is a multiple root (eigenvalue) of the characteristic equation, Pisarenko suggests taking successively smaller (principal) submatrices of $\underline{C}$ until ν becomes an isolated root (i.e., an eigenvalue with multiplicity one). As soon as ν is isolated, the eigenvector corresponding to is determined and the (complex) roots $\{z_j\}$ of the "null eigenvector polynomial" are computed. These roots are distinct and always lie <u>on</u> the unit circle (in the complex plane). As usual, the angular frequencies $\{\alpha_j\}$ are easily extracted from the roots (i.e., $\alpha_j = \arg z_j$). An equivalent, and probably more efficient, procedure is obtained by applying the "modern" Levinson recursion to the signal covariance matrix $\underline{C} - \nu \underline{I}$. Note that the <u>singular</u> signal matrix is obtained from $\underline{C}$ by applying the (eigenvalue) noise power cancellation idea discussed in Chapter II. At some point in the recursion, the computed prediction error will be zero and, at this point, the error filter weights provide the coefficients for the null eigenvector polynomial.

The representation in (4-38) for a correlation matrix suggests that the Pisaranko model is a special case of the general problem considered in Chapter II. In a direction finding context, the Pisarenko model is only valid when an ideal linear array receives incoherent signals. In this case, the sample covariance matrix $\hat{\underline{R}}(K)$ approaches the true correlation matrix $\underline{C}$ in (4-38) for

sufficiently large K. However, the remarkable feature of the Pisarenko decomposition is that it remains valid, in an _algebraic_ sense, for any correlation estimate. Thus, for finite K one first constructs a correlation estimate $\hat{\underline{C}}$ from $\hat{\underline{R}}(K)$ using one of the techniques discussed in Chapter 4, Section D. The Pisarenko method then generates estimates $\hat{\underline{\alpha}}$, $\hat{\underline{v}}$, and (diagonal) $\hat{\underline{\Psi}}$ such that

$$\hat{\underline{C}} = \underline{V}(\hat{\underline{\alpha}}) \; \hat{\underline{\Psi}} \; \underline{V}^H(\hat{\underline{\alpha}}) + \hat{v} \; \underline{I} \qquad (4\text{-}39)$$

Given $\hat{\underline{C}}$, Pisaranko's procedure or the modified Levinson recursion discussed above may be used to determine $\hat{v}$ and $\hat{\underline{\alpha}}$. Given these estimates, the remaining power estimates follow easily, i.e.,

$$\hat{\underline{\Psi}} = \hat{\underline{V}}^+ (\hat{\underline{C}} - \hat{v} \; \underline{I}) \; \hat{\underline{V}}^{+H} \qquad (4\text{-}40)$$

where $\hat{\underline{V}}^+$ represents the pseudo-inverse of $\underline{V}(\hat{\alpha})$.

Interestingly, a generalized version of the Pisarenko method surfaces from adaptive array theory if the usual linear constraint is replaced by a (quadratic) constraint on the _norm_ of the array weights. Thus, the problem

$$\sigma^2 = \min_{\underline{w}} \; \underline{w}^H \; \underline{R} \; \underline{w}$$

subject to $|\underline{w}|^2 = 1$ leads to array weights given by the eigenvector(s) of the smallest eigenvalue of $\underline{R}$.

Multiple Signal Classification (MUSIC) Algorithm

In [30], Schmidt proposed the MUSIC algorithm. In terms of our notation (Chapter II), Schmidt's direction-finding algorithm is

$$\{\hat{\underline{\alpha}}_i\} \; \Longleftarrow \; \underset{\alpha}{\text{maxima}} \; \frac{|\underline{v}(\alpha)|^2}{\underline{v}^H(\alpha) \underline{E}_N \underline{E}_N^+ \underline{v}(\alpha)} \qquad (4\text{-}41)$$

for any linear array. Of course, the idea behind MUSIC is that the denominator of (4-41) (theoretically) approaches zero for α sufficiently near any true angle α_i. Given exact covariance information, the MUSIC algorithm does indeed produce (very) large peaks. However, when the "null" eigenvectors are estimated from a reasonable finite number of snapshots, the MUSIC "spectral" peaks are not particularly sharp. The accuracy of the MUSIC angle estimates is not appreciably better than other more computationally efficient super-resolution algorithms; moreover, its resolution (detection) capability may be significantly worse (see section V.B.2).

If no spatial averaging were employed, MUSIC would fail completely in the coherent case. Of course, the same is true for MLM, MEM, etc. In the simulation results presented subsequently, the minimal amount of spatial smoothing needed to guarantee perfect asymptotic performance was used in all cases.

Eigen Assisted Autoregressive Analysis (EAR)

In Chapter III, it was shown that the multiple snapshot maximum likelihood (ML) angle processor behaves asymptotically like a single snapshot processor operating on an eigenvector of the sample covariance matrix. Since the exact ML processor is difficult to implement, even in the single snapshot case, an approximation based on a more practical processor is proposed. The increased sensitivity of the AR rooting method discussed in section D suggests that ARM be applied to the principal components of the sample covariance matrix (i.e., the eigenvectors with large eigenvalues). Thus, one employs the modified covariance algorithm to first smooth the signal space projection matrix $\underline{E}_S\underline{E}_S^+$ where $\underline{E}_S$ is a matrix constructed from the principal components of the sample covariance matrix[*]. This procedure is similar in spirit to noise cancellation approaches which attempt to remove the noise power term along the main diagonal of the covariance matrix. The angle estimates are then extracted from the smoothed covariance "estimate" exactly as in the ARM algorithm

---

[*]We may interpret $\underline{E}_S\underline{E}_S^+$ as a processed covariance matrix (estimate) which contains only the signal directional information (see Chapter II).

4-38

described previously. The resulting algorithm is called the Eigenvector Autoregressive Root method (EAR).

V. PERFORMANCE ASSESSMENT

A.   Assessment Approach

In Chapter III, we obtained fundamental theoretical limits on angular accuracy using the Cramer-Rao (CR) bound. Maximum likelihood (ML) angle estimates which asymptotically approach the CR bound were investigated, but algorithms for actually generating the ML estimates were (later) found to be computationally unattractive.   Subsequently, a number of more practical approaches to the angle estimation problem were described in Chapter IV.

In this chapter, we present performance results obtained via Monte Carlo and deterministic simulations for several promising super-resolution schemes. At high SNR, the angular accuracy of these methods has been found to be in reasonably good agreement with predicted values derived from the Cramer-Rao theory.

One of the important issues considered is the performance improvement which may be obtained by time averaging intermediate statistics (especially, the covariance) and making an angle estimate on the averaged statistic as opposed to averaging angle estimates. One naturally expects an improvement in performance as the number of snapshots available for averaging increases. Our simulation results confirm this conjecture, although in some cases the improvement may be more or less pronounced than is predicted by the usual $1/\sqrt{K}$ rule-of-thumb, where K is the number of independent observations.

One might well question the motivation for multiple-snapshot processors, if averaging the output of a single-snapshot processor works just as well. In the first place, a multiple-snapshot processor is probably more efficient and will invariably have a lower detection threshold. Moreover, the angle estimation performance immediately above the threshold may lie closer to the CR bound.

While the preceding remarks are quite general and hence somewhat imprecise, they serve to emphasize the point that angular accuracy at very high SNR is not necessarily the best criterion for assessing the "resolution" performance of direction-finding algorithms. In fact, the ability of a processor to

5-1

identify the correct number of signals present, especially at low SNR, may be a much more fundamental issue than angular accuracy. Power estimates are frequently required and, in some cases (e.g. multipath environment characterization), estimates of relative signal phase may be desired. Finally, one has hopes that the problems of target splitting and spurious peaks that have recently received so much attention in the spectral estimation community [2] will not degrade system performance at high SNR. Taken collectively, the characteristics just described constitute a multi-faceted but somewhat nebulous definition of resolution. In this chapter, these criterion are refined and statistically quantified in a manner that will support an objective assessment of the capability of any high-resolution, direction-finding technique.

1. Decision Theoretic Assessment

The performance assessment approach developed for assessing high-resolution, direction-finding algorithms is illustrated in Fig. 5-1. The algorithm under test supplies data to the performance assessment (sub) program in the form of signal reports. Each such report consists of a list of paired direction-of-arrival (DOA) and signal-to-noise ratio (SNR) estimates. This list is compared to the true DOA/SNR values. A suitable measure of error is achieved with the help of an assignment algorithm which first associates the signal estimates witn the signals actually present. The assignment strategy minimizes the sum of the angular miss distances subject to the provision that no assignment will be made that results in an angle error greater than a beamwidth or a power error greater than 10 dB. In the event that more signals are reported than can be assigned, an appropriate number of false alarms is recorded. Similarly, a miss counter is incremented for each true signal that cannot be associated with any of the signal reports. Of course, the angle and power errors for each successful match are recorded. After a pre-determined number of independent trials, statistical estimates of the false alarm rate, miss probabilities, angle and power bias, and angle standard deviation are computed from the Monte Carlo results. The Monte Carlo statistics may be plotted and/or compared to pre-determined limits that effectively constitute a

TARGET* REPORTS

*from angle of arrival
 estimation algorithms

ASSIGNMENT ALGORITHM

TRUE TARGET DATA

- MINIMIZE MISS DISTANCES
- ANGLE ERROR < 1 Bw
- POWER ERROR < 10 dB

ERROR METRICS

- MISSES
- FALSE ALARMS
- ANGLE ERROR
- POWER ERROR

MONTE CARLO
AND
PARAMETER VARIATION
LOOPS

STATISTICS PLOTS
(vs SNR)

RESOLUTION CRITERIA:

	VIOLATION CODE
PROB. OF MISS < 10%	▮
FALSE ALARM RATE < 10%	▦
|MEAN POWER ERROR| < 3 dB	▤
|ANGLE BIAS| < BEAMWIDTH / 30	▨
ANGLE STD. DEV. < BEAMWIDTH/30	▨

RESOLUTION BAR GRAPH
(Emitter separation vs SNR)

Fig. 5-1. Performance assessment approach.

statistical definition of resolution. Naturally, the resolution quantifiers
are somewhat subjective and may be mod*fied to fit the requirements for a
specific application.

In order to illustrate our approach, the criteria shown in Fig. 5-1 will
be used to assess the performance of four of the algorithms discussed in the
previous chapter. The details of the simulation are discussed in Section B-1,
and results are presented in Section B-2.

2. Asymptotic Performance

The Monte Carlo decision theoretic approach discussed above is quite
comprehensive, but requires significant amounts of computer time to determine
performance trade-offs.   Consequently, we augmented this Monte Carlo
simulation with an alternative simulation model in which the observable is the
ensemble overage covariance matrix discussed in Chapter II.   This particular
observable yields meaningful results only for the estimators (e.g., MEM, MLM,
ARM) which do not take advantage of the asymptotic covariance properties.

3. Assessment of Array Errors

The Monte Carlo and deterministic simulations outlined above examine the
impact of additive noise (typically, from the receiver front ends) on the
various algorithms.   Since this error source is analogous to the additive
white Gaussian noise typically encountered in time-series spectrum estimation,
there has been considerable attention to its effects on high-resolution tech-
niques [2].   However, there are several other error components which also
degrade performance that are not easily incorporated into the additive noise
representation. Principal among these are:

(1) A/D conversion noise

(2) channel gain and phase errors

(3) antenna element pattern differences

Below we discuss each of these and indicate how they have been addressed in
the context of the present study.

The effects of round-off errors in the signal A/D conversion process will depend on the quantity converted. Conversion of in-phase (I) and quadrature (Q) components gives rise to an additive error which in most cases can be regarded as analogous to front-end noise. The resulting noise is small when the input signal levels are high, but can be significant at low signal levels if there is a wide dynamic range of received signal strengths. Consequently, a number of practical systems (including that used for the experiments discussed in Section D) digitize log amplitude and phase so as to obtain a roughly constant quantization error characteristic over the full range of signal levels. No general theory of the relative effects of amplitude versus phase errors has been achieved for the angle estimators of greatest interest.* Consequently, the amplitude/phase quantization effects have been examined both by computer simulations and (implicitly) in the field tests.

Careful channel matching in both amplitude and phase is essential since errors in the digitized wavefront may be interpreted by the adaptive algorithms as corresponding to additional signal sources and/or cause errors in the estimated position of the direct signal. Similar effects arise with conventional systems; however, the high resolution algorithms all employ highly nonlinear operations on the received data. Thus, the effects of channel mismatch may be far more pronounced with the high-resolution algorithms.

As is the case of quantization errors, no analytical theory has been developed to date to predict the degradation as a function of the amplitude and phase errors. Simulation studies were carried out and considerable attention was paid to reducing and quantifying the degree of channel mismatch in the field experiments (see Chapter VI).

Differences in the various antenna element patterns (due to element differences and/or mutual coupling) produce angle-dependent amplitude and phase errors similar to those due to poor channel equalization. In the experimental data, particular care was taken to minimize these effects by:

---

*It was shown in the Doppler MLS studies that amplitude errors will not, to first order, affect the null position for a real (Δ/Σ) monopulse processor [29].

(1) using elements which were as physically identical as possible,

(2) using relatively large element spacings,

(3) adjusting of the array physical configuration to minimize coupling effects.

The details of the minimization process are presented in Chapter VI and Appendix A.

To summarize, conventional antenna array and receiving techniques result in measurement amplitude and/or phase errors which cannot easily be incorporated into the additive noise modelling discussed in the preceding sections. Certain of these errors (in particular, A/D conversion round-off errors) could be reduced by time averaging. However, many of the other errors represent unknown biases which must be minimized by careful system design. In the absence of analytical theory to predict the effects of these errors on the various algorithms, we have been forced to examine the effects via computer simulations and field experiments.

Appendix C describes simulations designed to provide insight into the tolerable amplitude and phase errors for the more promising algorithms. The specific model adopted was to assume independent log amplitude and phase errors which were each uniformly distributed over an interval centered at zero error. This particular model is viewed as quite appropriate for A/D quantization and reasonable for channel equalization or element gain errors (since the system calibration process should remove any tails of the equalization process). Element pattern differences as a function of azimuth and/or elevation may have somewhat larger tails, but should at least be independent from element to element. Mutual coupling effects between the elements can yield correlated phase and amplitude errors (see Chapter VI), but development of a detailed model was not possible in the context of the present study.

5-6

B.   Decision Theoretic (Monte Carlo) Simulations

1.   Simulation Description

Figure 5-2 is a conceptual flow diagram of the (sub) program that generates simulation data for the performance assessment algorithm of Fig. 5-1. This particular direction-finding simulation has five essentially distinct parts:

(1)   signal/array model
(2)   sample covariance generation
(3)   covariance (data) processing options
(4)   angle estimation options
(5)   power estimation/threshold test

In addition, three nested loops are provided in order to conduct Monte Carlo and/or parameter variation studies. The basic output of this program is a collection of target (i.e., signal) reports. Each target report is a list of paired angle/power estimates.

Currently, the simulation program handles the special case of two signals received by an ideal linear array (i.e., a linear array with uniformly spaced, isotropic elements). Up to 32 array elements may be specified. The relative amplitude and phase of the signals and/or the array signal-to-noise ratio may be varied. The signals may be coherent, incoherent, or partially coherent. Mutual coupling effects have not been included, and the possibility of ambiguous direction estimates is ignored.

Sample covariance matrices consistent with the selected signal parameters are generated in accordance with a complex Gaussian model for the observed data. The (receiver) noise components contributed by the individual array elements are assumed to be statistically independent and identically distributed random variables. The signal-in-space is modelled as a two-dimensional, zero-mean complex Gaussian vector with a (2x2) covariance matrix consistent with the specified level of signal coherence.
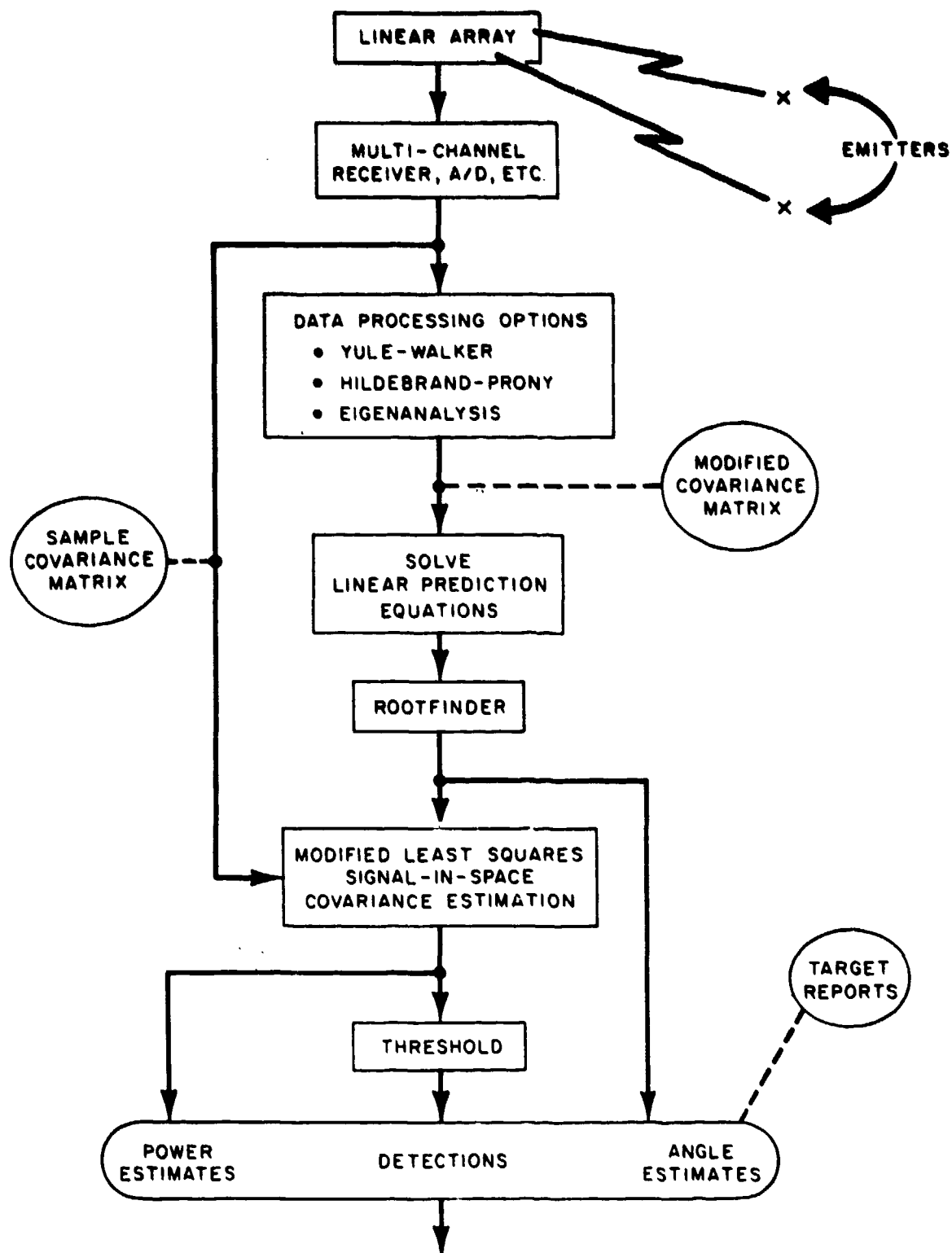
Fig. 5-2. Generation of plane wave "Target Reports" in Monte Carlo simulation.

The number of snapshots (observations) used to construct each sample covariance matrix is a simulation parameter. In order to avoid excessively high computation costs when the number of snapshots is large, an efficient simulation technique based on the complex Wishart distribution [72] has been developed. In this technique, the array data $\underline{r}(k)$ are described statistically by

$$\underline{r}(k) = \underline{L} \, \underline{w}(k) \tag{5-1}$$

where $\underline{w}(k)$ is a white "innovations" process normalized to unit variance. Note that $\underline{L}$ is the lower triangular matrix obtained from the standard Cholesky decomposition of the covariance $\underline{R}$ of the observed data, i.e.,

$$\underline{R} = \underline{L} \, \underline{L}^H \quad . \tag{5-2}$$

Clearly, any sample covariance matrix $\underline{\hat{R}}(K)$ may be written as

$$\underline{\hat{R}}(K) = \underline{L} \, \underline{\hat{W}}(K)\underline{L}^H \tag{5-3}$$

where $\underline{\hat{W}}(K)$ is an equivalent "sample" covariance for the innovations process.

The power of the representation in (5-3) stems from the fact that a Cholesky-type representation for $\underline{\hat{W}}(K)$ can be generated very efficiently. Following Goodman [72], we write

$$\underline{\hat{W}}(K) = \underline{\hat{T}}^H(K)\underline{\hat{T}}(K) \tag{5-4}$$

where $\underline{\hat{T}}(K)$ is an upper triangular matrix. From the probability density function for $\underline{\hat{T}}$ ([72], Eq. (1.9)), it can be shown that the (non-zero) elements of $\underline{\hat{T}}$ are zero-mean, (circular) complex Gaussian variates; a number of well known techniques (e.g., Box-Mullen) are available that generate Gaussian variates efficiently. The $\underline{\hat{T}}$ are jointly statistically independent in the special case $E\{\underline{\hat{W}}\} = \nu\underline{I}$ (where ν is any positive real number). In this case, the diagonal elements of $\underline{\hat{T}}$ can be conveniently constructed by taking square roots

of statistically independent gamma variates. The IMSL (Fortran) routine GGMAR [24] generates gamma variates very efficiently using the (exact) method of acceptance/ rejection testing [81].

The covariance processing options included in the simulation break down into three general categories. The fundamental covariance spatial averaging techniques were analyzed in some detail in Chapter II. Correlation (lag) estimates were briefly reviewed in Chapter IV, Section D. Eigenvector methods are discussed in Chapter IV, Section F. Loosely speaking, the results of covariance processing may be viewed as a "modified" covariance estimate. In the case of correlation estimates, the "modified" covariance estimate is a Toeplitz matrix. When eigenvector methods are employed, the result may generally be viewed as a projection matrix which _may_ be "smoothed" via any of the aforementioned techniques.

Angle estimates may be extracted from an arbitrary covariance matrix $\underline{R}$ by one of three basic techniques. The beam scanning (BS) method is analogous to conventional spectral analysis and amounts to computing the quadratic form

$$\hat{\Psi}_{BS}(\alpha) = \underline{v}^H(\alpha)\underline{R}\,\underline{v}(\alpha) \tag{5-5}$$

where $\underline{v}(\alpha)$ is the vector of (complex) array weights that "steers" the array (beam) to the angle α. In the special case of an ideal linear array, Eq. (5-5) may be implemented by incoherently averaging FFT(s) of vectors derived from any convenient "square root" of $\underline{R}$ (e.g., Cholesky, Hermitian, etc.). This approach is used in the simulation to compute a spatially-averaged "MUZAK" spectrum

$$\hat{\Psi}_{MUZAK}(\alpha) = |\underline{v}^H(\alpha)\,\hat{\underline{E}}_S|^2 \tag{5-6}$$

where $\hat{\underline{E}}_S$ is the matrix of principal components (see Chapter IV, Section F) of the spatially averaged covariance estimate $_2\hat{\underline{R}}_M(K)$. Note that the MUZAK spectrum can be interpreted as a _modified_ scanning beam spectrum.

When the data from an isotropic array are not spatially averaged, the MUZAK spectrum Eq. (5-6) is, for all intents and purposes, equivalent to the MUSIC spectrum Eq. (4-4). The equivalence follows by invoking the resolution of the identity (2.461) and the fact that $|\underline{v}(\alpha)|^2 \equiv 1$ for a (normalized) isotropic array.

The maximum likelihood method for computing spectral estimates was discussed in Chapter IV, Section C. The MLM spectral estimate

$$\hat{\underline{\psi}}_{MLM}(\alpha) = \left(\underline{v}^H(\alpha) \underline{R}'^{-1} \underline{v}(\alpha) \right)^{-1} \tag{5-2}$$

is not currently an option in the Monte Carlo performance assessment simulation, but is used in the deterministic simulations.

In the maximum entropy method(s), the Wiener equation

$$\underline{R}' \underline{w} = (\underline{1} \; 0 \; \ldots \; 0)^T$$

is "solved" for the "optimum" error filter weights $\hat{\underline{w}}$, and the MEM spectral estimate is computed via

$$\hat{\underline{\psi}}_{MEM}(\alpha) = \frac{constant}{|\underline{v}^H(\alpha)\hat{\underline{w}}|^2} \; . \tag{5-8}$$

The three fundamental techniques discussed above generate (continuous) power spectral density estimates. In contrast, the signal spectrum that arises in the direction-finding problem posed in Chapter II consists of a finite number of spectral "lines" (impulses). The standard way to proceed in these circumstances is to identify the spectral "line" frequencies from the local maxima of the power spectral density function. In the case of an array with identical and uniformly spaced elements, the "line" estimates may be derived directly from $\hat{\underline{w}}$ via the autoregressive root method (ARM) discussed in Chapter IV, Section D.

Given any set (vector) of angle estimates $\hat{\underline{\alpha}}$, the signal-in-space covariance matrix $\underline{P}$, introduced in Chapter II, may be estimated in a manner analo-

gous to the technique incorporated in Pisarenko's method (see Chapter IV, Section F), i.e.,

$$\hat{\underline{P}} = \hat{\underline{V}}^{+H}(\hat{\underline{R}} - \nu\underline{I})\hat{\underline{V}}^{+} \tag{5-9}$$

where $\hat{\underline{V}}^{+}$ is the pseudo-inverse of $\underline{V}(\hat{\underline{\alpha}})$, $\hat{\underline{R}}$ is the original covariance estimate, and ν is the true noise power level (assumed to be measurable and hence known). The diagonal elements of $\hat{\underline{P}}$ provide estimates of the power received from the various signal directions. In order to reduce the number of spurious angle estimates, each power estimate is compared to a threshold (usually related to the noise level). When a power estimate is below the (signal) threshold, that power estimate and the corresponding angle estimate are discarded. After the threshold test, the remaining angle/power estimates constitute a target report.

2. Monte Carlo Simulation Results

The relative performance of MEM, ARM, MUSIC, and EAR has been studied in a simulated specular multipath environment using the decision theoretic Monte Carlo package discussed above. Preliminary work had suggested that the performance of these (in fact, most) methods is poorest when the relative phase between the direct and indirect signal is 0° and is best at 90° relative phase. In the simulation runs reported here, only these two extreme cases are considered. The angular separation between the two signals was systematically varied between 0.1 and 0.8 beamwidth in 3-dB steps. An ideal array with nine uniformly spaced elements was postulated. The relative amplitude of the two signals was 1 dB and the relative phase is defined with respect to the center element.

In most cases, the array signal-to-noise ratio (per snapshot) was varied between 10 and 50 dB in 5-dB increments. The benefits of time-averaging were examined by considering sample sizes of 1, 10, and 100 snapshots. In order to determine performance thresholds with reasonable granularity, and without undue expense, all of the simulation statistics generated as a function of SNR were linearly interpolated. A study of the simulation statistics presented in Figs. 5-3 through 5-6 indicate that no significant errors are introduced by this simple interpolating stratagem.
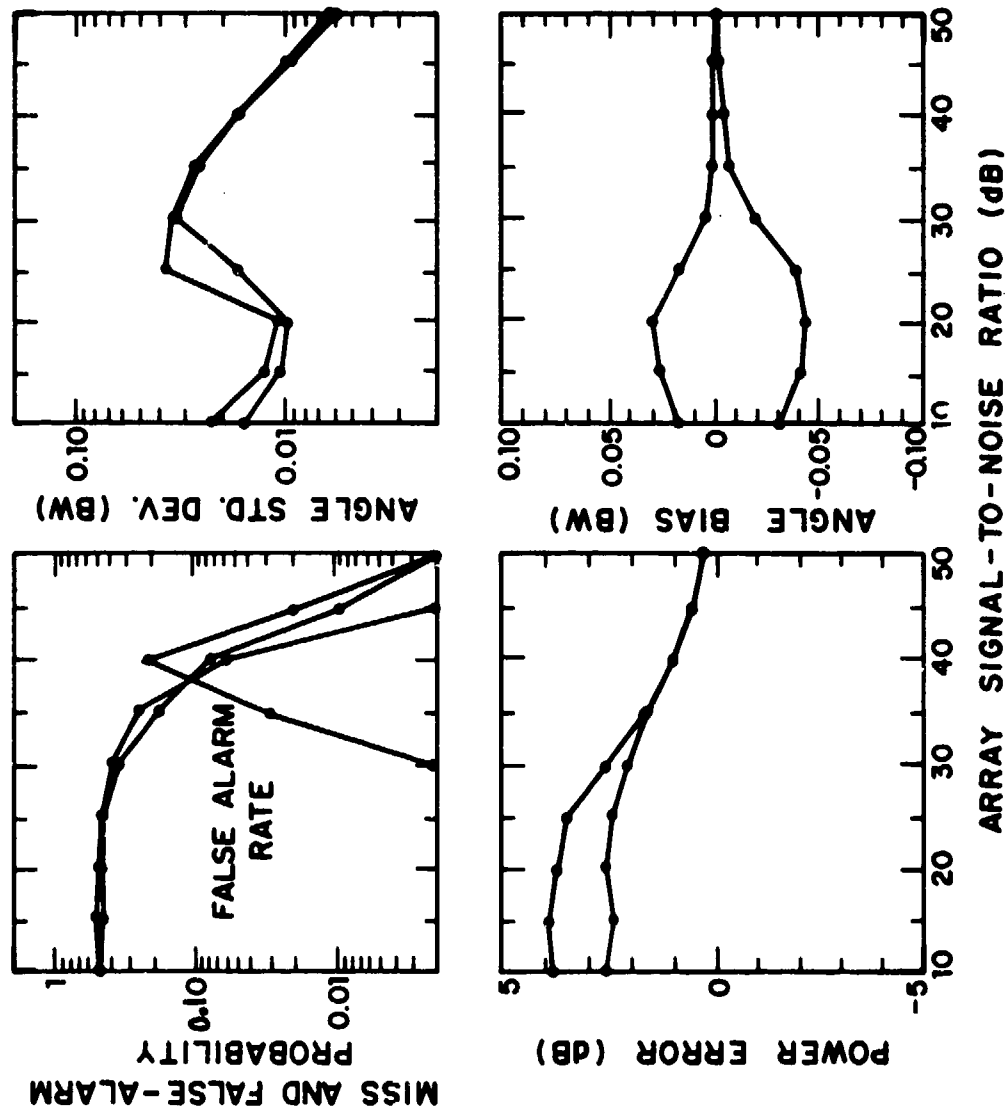
TWO COHERENT SIGNALS
(relative amplitude = 1 dB)

● RELATIVE RF PHASE = 90°
● ANGULAR SEPARATION = 0.1 BW
● NO OF SNAPSHOTS = 10

ANGLE STD. DEV. (BW)
0.10
0.01

ANGLE BIAS (BW)
0.10
0.05
0
-0.05
-0.10

MISS AND FALSE-ALARM PROBABILITY
1
0.10
0.01

FALSE ALARM RATE

POWER ERROR (dB)
5
0
-5

ARRAY SIGNAL-TO-NOISE RATIO (dB)
10 20 30 40 50

110475-N,A,B,C

Fig. 5-3. MEM simulation statistics.
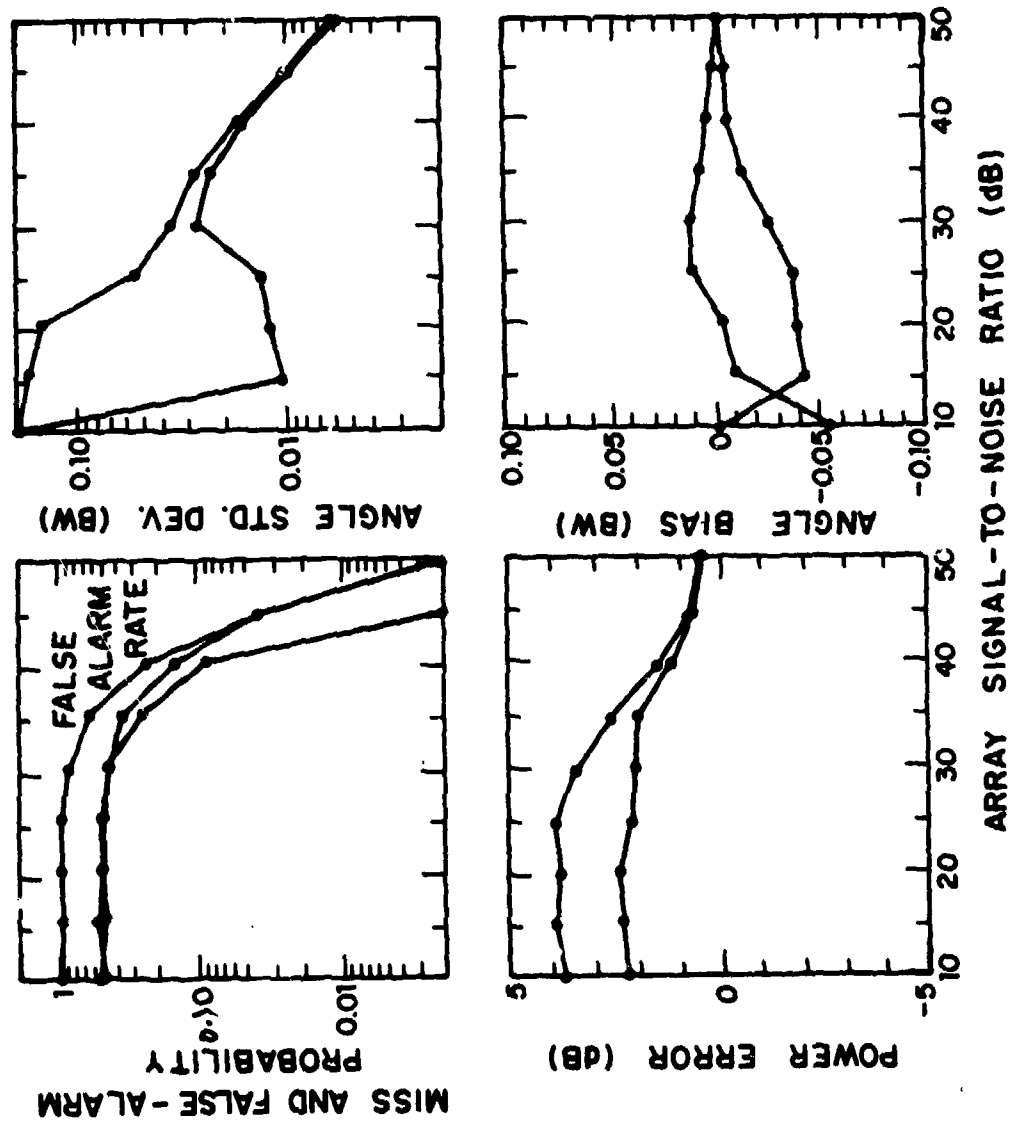
5-13

TWO COHERENT SIGNALS
(relative amplitude = 1 dB)

- RELATIVE RF PHASE = 90°
- ANGULAR SEPARATION = 0.1 BW
- NO. OF SNAPSHOTS = 10

Fig. 5-4. MUSIC simulation statistics.
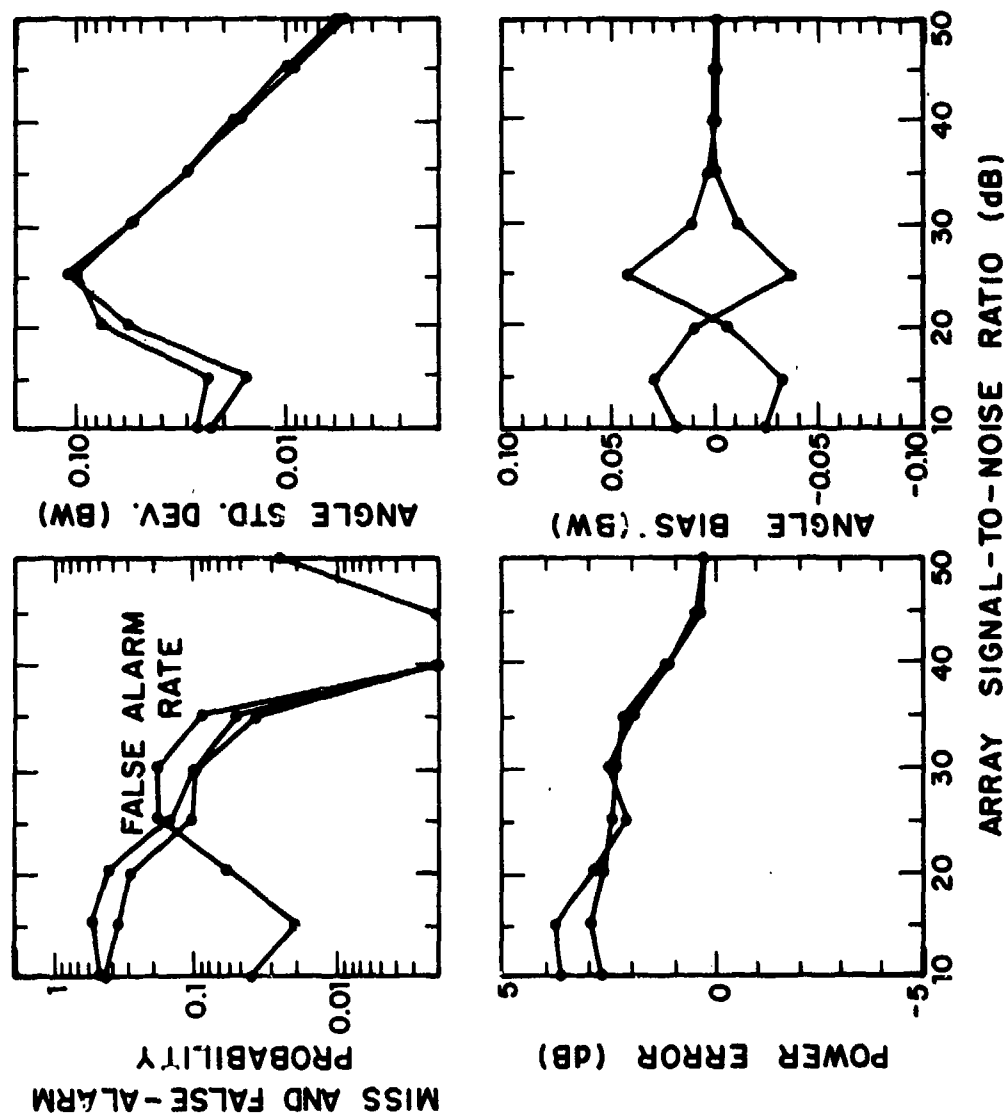
5-14

Fig. 5-5. ARM simulation statistics.
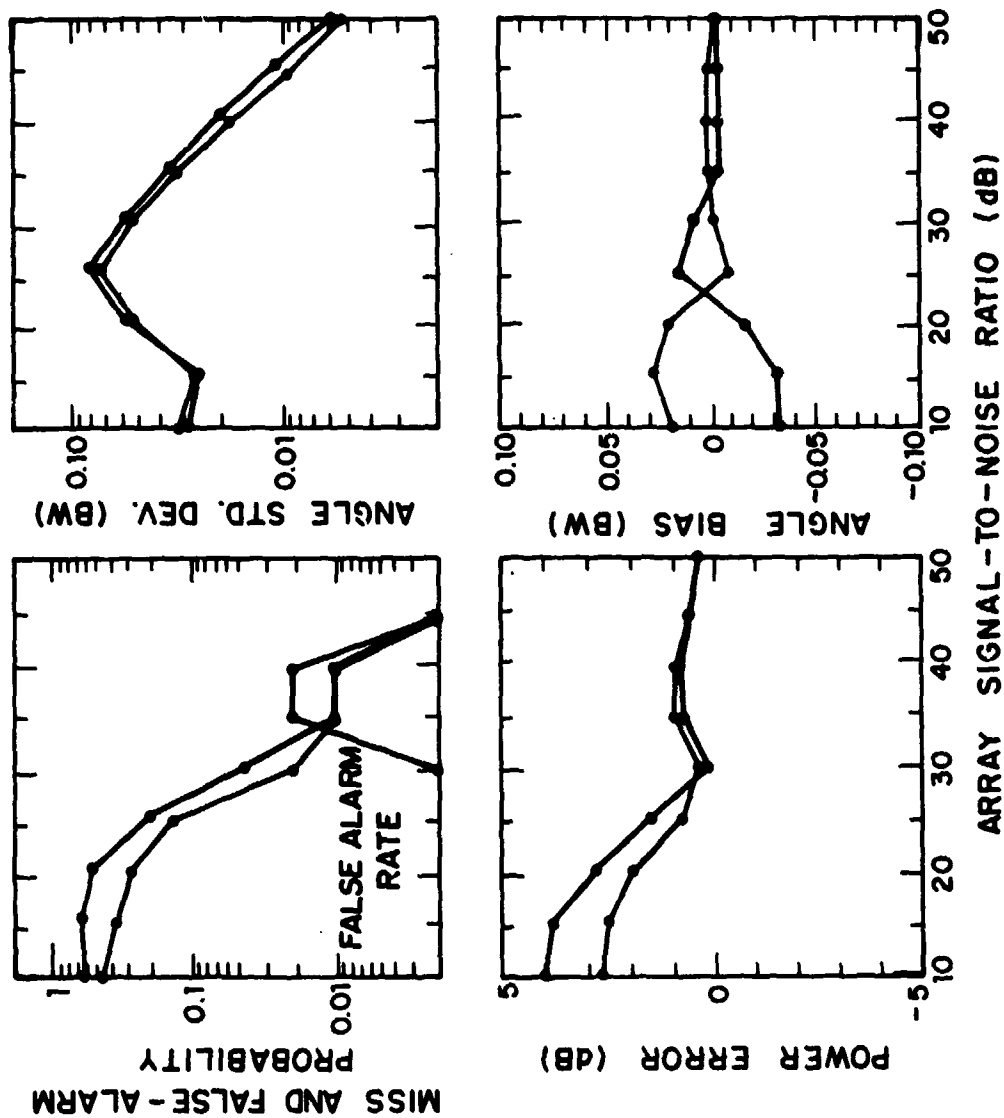
TWO COHERENT SIGNALS
(relative amplitude = 1 dB)

- RELATIVE RF PHASE = 90°
- ANGULAR SEPARATION = 0.1 BW
- NO OF SNAPSHOTS = 10

110477-N,A,B,C

TWO COHERENT SIGNALS
(relative amplitude = 1 dB)

● RELATIVE RF PHASE = 90°

● ANGULAR SEPARATION = 0.1 BW

● NO OF SNAPSHOTS = 10

110478-N,A,B,C

Fig. 5-6. EAR simulation statistics.

5-16

Using the decision theoretic criteria in Fig. 5-1, the resolution capabilities of MEM, ARM, MUSIC, and EAR have been evaluated while varying the number of snapshots as discussed above. The results for the favorable 90° relative phase case are presented in Figs. 5-7 through 5-9. The minimum array SNR required to satisfy all of our resolution criteria has been tabulated in Table 5-1. These data are taken directly from the bar graphs in Figs. 5-7 through 5-9. A similar tabulation is provided in Table 5-2 for the worst case relative phase (0°).

The simulation results suggest that EAR provides the best angular resolution when a large number of snapshots are available. If only one snapshot is processed, angular resolution is rather severely limited by the array SNR and all of the methods perform about the same. Given a moderate number of snapshots (i.e., approximately the same as the number of antenna elements), the signal processing performance improves substantially and, as in the single-look case, all of the methods fare equally well.

C. Simulation Results with Deterministic Observation Models

The asymptotic performance of various estimators as the number of time observations (K) approaches infinity is of interest in that the performance of many estimators closely approximates their asymptotic performance for relatively small numbers of time samples. This approach is also useful for providing performance parametric studies with a smaller computer time expenditure.

The eigenanalysis techniques (i.e., MUSIC and EAR) will generally work perfectly in the asymptotic limit provided that sufficient (i.e., some) spatial averaging is accomplished and the number of signals to be resolved is sufficiently less than the number of array elements (as indicated in Chapter II). Thus, the attention here was focused on the other estimators (e.g., MEM, ARM, MLM) which do not explicitly take advantage of an infinite number of time observations.

5-17
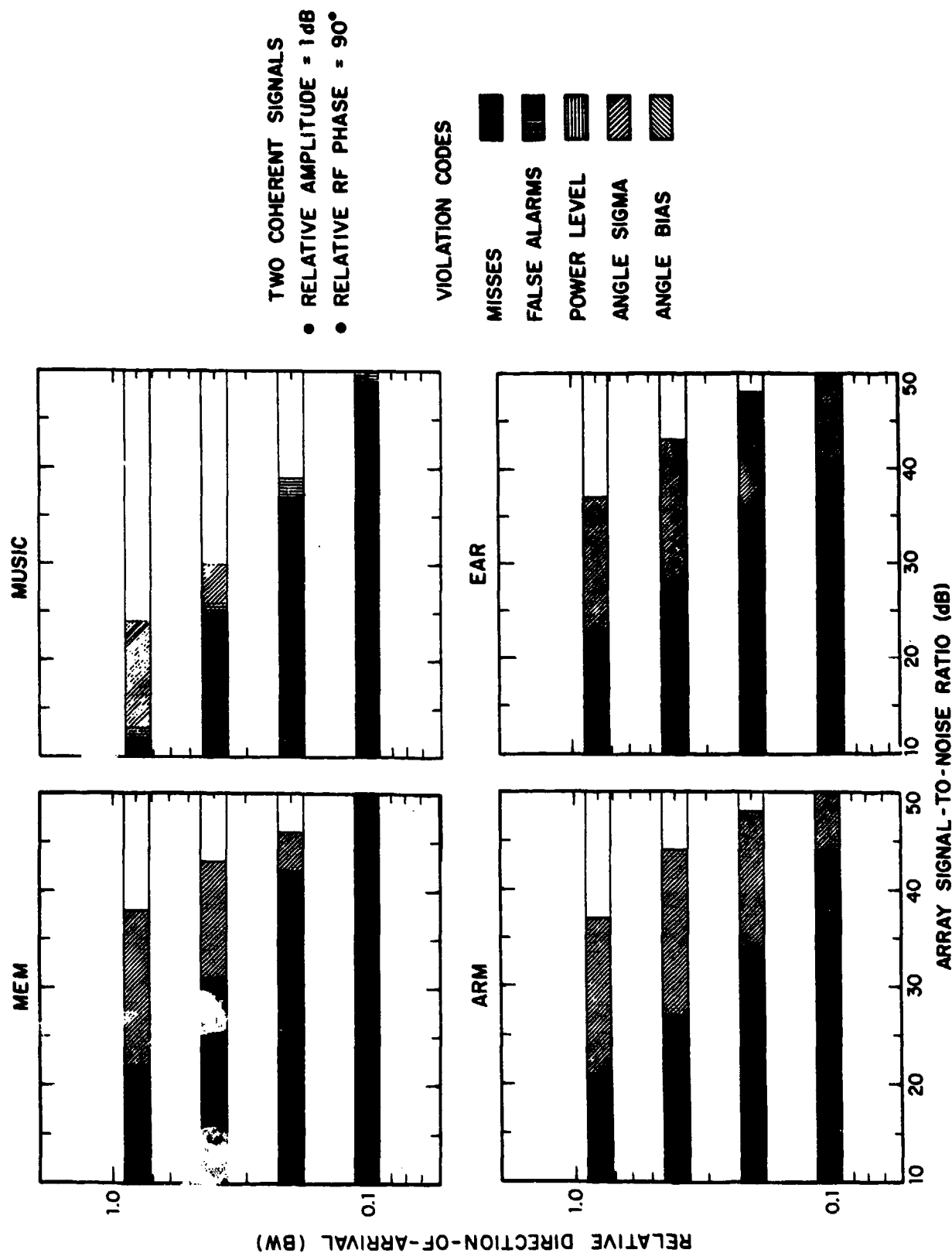
Fig. 5-7. Comparative coherent signal resolution performance: 1 snapshot.
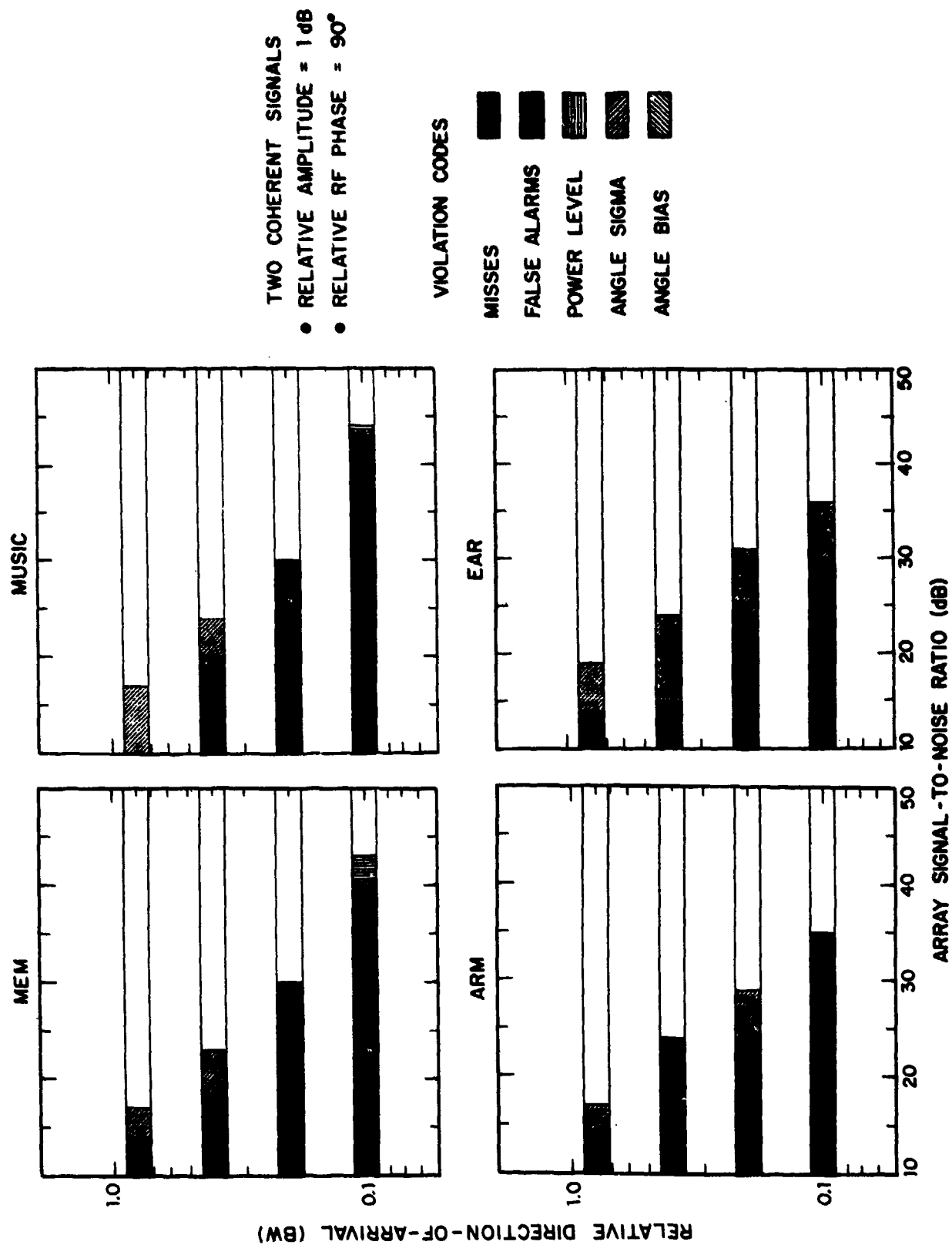
Fig. 5-8. Comparative coherent signal resolution performance: 10 snapshots.
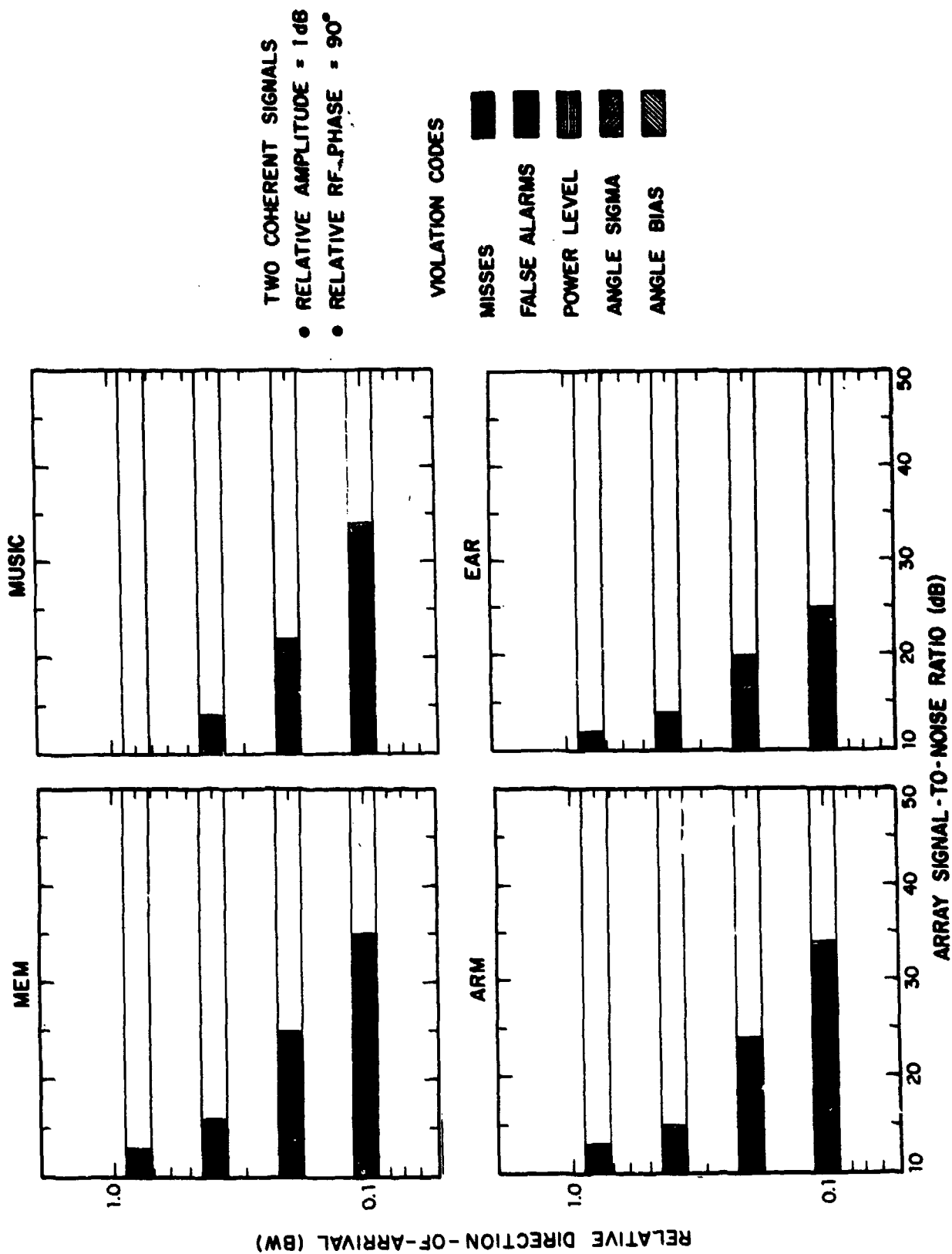
Fig. 5-9. Comparative coherent signal resolution performance: 100 snapshots.

5-20

TABLE 5-1

MINIMUM ARRAY SNR TO RESOLVE TWO SIGNALS 90° OUT-OF-PHASE
(RELATIVE AMPLITUDE = 1 DB)

METHOD \ ANGULAR SEPARATION	0.2 BW	0.4 BW	0.8 BW
MUSIC	49 dB	40 dB	34 dB
MEM	46 dB	43 dB	38 dB
EAR	48 dB	43 dB	37 dB
ARM	48 dB	44 dB	37 dB

a) 1 Look

METHOD \ ANGULAR SEPARATION	0.1 BW	0.2 BW	0.4 BW	0.8 BW
ARM	35 dB	29 dB	24 dB	17 dB
EAR	36 dB	31 dB	24 dB	19 dB
MEM	43 dB	30 dB	23 dB	17 dB
MUSIC	44 dB	30 dB	24 dB	17 dB

b) 10 Looks

METHOD \ ANGULAR SEPARATION	0.1 BW	0.2 BW	0.4 BW	0.8 BW
EAR	25 dB	20 dB	14 dB	12 dB
MUSIC	34 dB	22 dB	14 dB	<10 dB
ARM	34 dB	24 dB	15 dB	13 dB
MEM	35 dB	25 dB	16 dB	13 dB

c) 100 Looks

5-21

TABLE 5-2

MINIMUM ARRAY SNR TO RESOLVE TWO IN-PHASE SIGNALS
(RELATIVE AMPLITUDE = 1 DB)

ANGULAR SEPARATION / METHOD	0.4 BW	0.8 BW
MEM	58 dB	43 dB
MUSIC	59 dB	44 dB
ARM	58 dB	47 dB
EAR	58 dB	47 dB

a) 1 Look

ANGULAR SEPARATION / METHOD	0.4 BW	0.8 BW
EAR	39 dB	26 dB
MUSIC	40 dB	25 dB
MEM	42 dB	29 dB
ARM	43 dB	29 dB

b) 10 Looks

ANGULAR SEPARATION / METHOD	0.2 BW	0.4 BW	0.8 BW
EAR	39 dB	27 dB	16 dB
MUSIC	48 dB	30 dB	17 dB
MEM	>50 dB	33 dB	14 dB
ARM	49 dB	35 dB	24 dB

c) 100 Looks

The observation model here is the ensemble average covariance $\underline{R} = \underline{V} \, P \, \underline{V} + \underline{N}$ which was introduced in Chapter II. We considered the two limiting cases for $\underline{P}$, namely:

(i) a singular $\underline{P}$ whereby $P_{nm} = P_n^2 \, \rho_{mn} e^{j\phi_{mn}}$ where ρ_{mn} is the relative amplitude of signal m to signal n and ϕ_{mn} is the relative rf phase at the array center (this would correspond to multipath with low scalloping rates), and

(ii) a diagonal $\underline{P}$ whereby $P_{mn} = p_m^2 \delta_{mn}$ corresponding to uncorrelated signal sources or multipath with high scalloping rates. The diagonal $\underline{P}$ (i.e., incoherent) case is believed to provide a lower bound on the coherent performance as well as being a benchmark by which our numerical results can be compared with those of other researchers.

The degree of covariance spatial smoothing used to obtain numerical results in these two cases was small. In the case of incoherent signals, all diagonals of R are equal (i.e., the covariance matrix is Toeplitz) and the best performance is typically obtained by using the largest possible order for a given estimator. With a singular P, we have found that only pairwise smoothing (i.e., a model order which is one less than that used for the incoherent case or, two subapertures) yielded the best results.

1. Two Plane Wave Cases

Figures 5-10 and 5-11 show the MEM, MLM and conventional beamsum spectra at fixed separation for incoherent signals. For those cases, the spectrum peaks generally fell within 0.05 beamwidth of the true angle provided that there was at least a 1-dB dip between adjacent peaks. Thus, the criteria

(a) proper number of peaks which are at least 3 dB above the noise level, and

(b) at least 1 dB dip between peaks

were used to define successful operation. By repeating such simulations at a variety of separation angles and SNR values, it was possible to define thresholds for successful performance as shown in Fig. 5-12. Also shown in Fig. 5-12 is the MLM tradeoff curve of Cox [3] which was shown earlier in Fig. 4-5. The good agreement between Cox's results and ours is viewed as confirming the
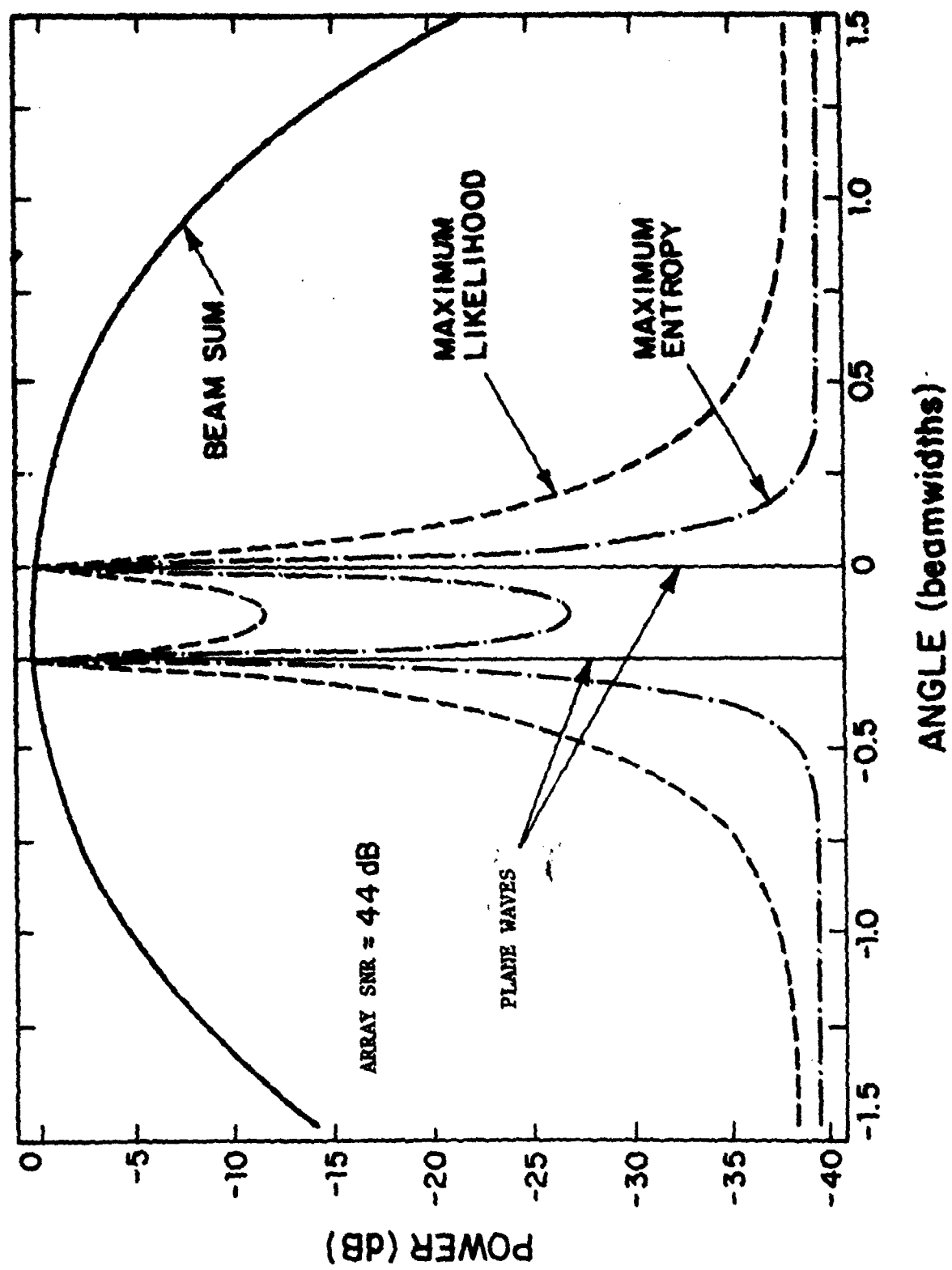
Fig. 5-10. Example of resolution of two equal power incoherent plane waves.

18-00-18308-1

5-24
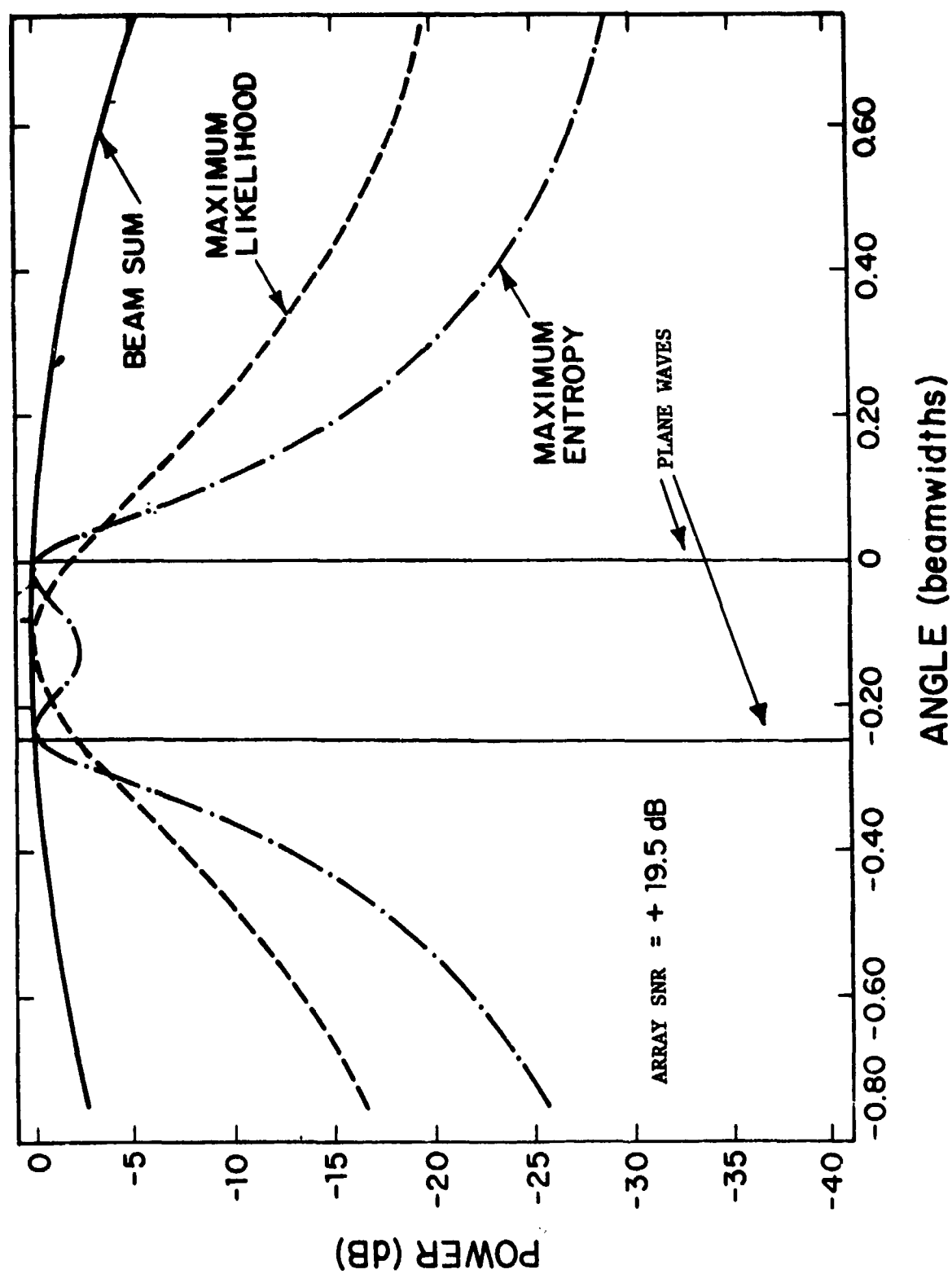
Fig. 5-11. Example of resolution of equal power incoherent plane waves.
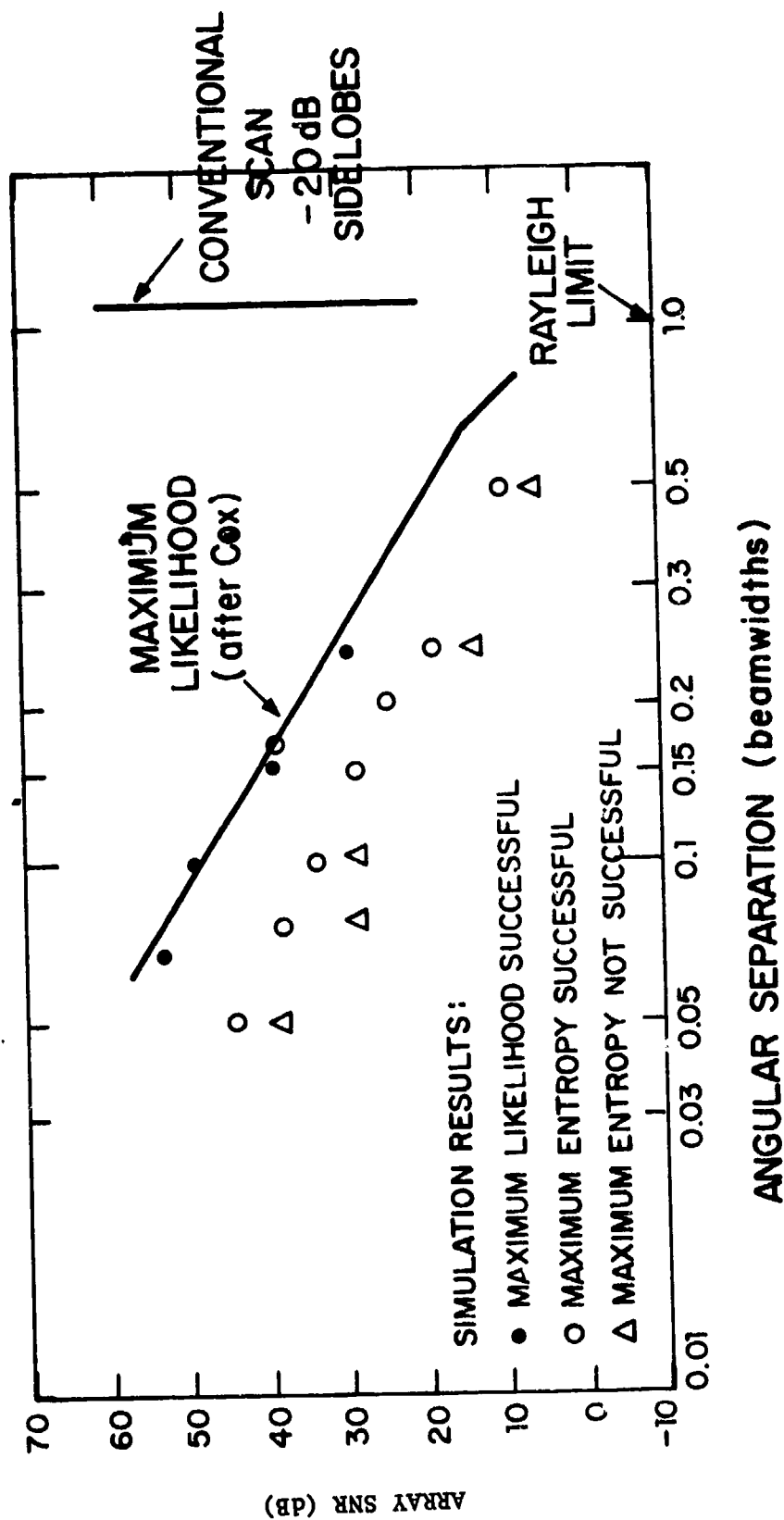
Fig. 5-12. Comparison of angular spectrum resolution limits for equal power incoherent plane waves.

5-26

accuracy of our numerical computations. The MEM results showed similar good agreement with the resolution curves obtained by Gabriel [1] using the Applebaum-Howells adaptive nulling array algorithm.

Based on Cox's result that the MLM resolution performance is very weakly dependent on relative signal levels provided that the SNR is that of the weaker signal, simulations were made for incoherent signals with relative power levels of 10 and 20 dB. The spectra in these case (Figs. 5-13 and 5-14) show the large power estimate biases[*] which can arise with the raw MEM spectra if an alternative power estimation technique such as Eq. (5-9) is not utilized. Figures 5-15 and 5-16 show the corresponding separation angle - SNR trade-offs for successful resolution angle estimation. We see that both the MLM and MEM performance is largely independent of relative power levels, whereas the beamsum resolution is significantly greater (i.e., a factor of 2) than the Rayleigh limit at large differences in relative power level.

Figure 5-17 compares the ARM technique with the other three techniques for incoherent signals which differ by 1 dB and 10 dB. We see that the ARM performance is substantially better at small separation angles. The performance improvement in this case with ARM can be explained in terms of the respective error filter zero locations as shown in Fig. 5-18. When the SNR is high, the principal zeros (i.e., those with largest modulus) lie on the unit circle at z values corresponding to the plane-wave arrival angles. As the SNR decreases, the zero corresponding to the weaker plane wave moves inside the unit circle. When the zero has moved inside the unit circle by an amount comparable to the z-plane displacement between the infinite SNR zero locations, the spectrum will typically show a single peak near the angle of the larger plane wave. However, the zero locations in the z plane still indicate the location of the infinite SNR zeros and hence can be used.

Figures 5-19 to 5-22 show the corresponding resolution thresholds of a 9-element array for two coherent plane waves differing in power by 4 dB with the largest possible model orders. Figures 5-23 to 5-25 are the corresponding

---

[*]Approximately a square-law suppression of the weaker signal.

5-27
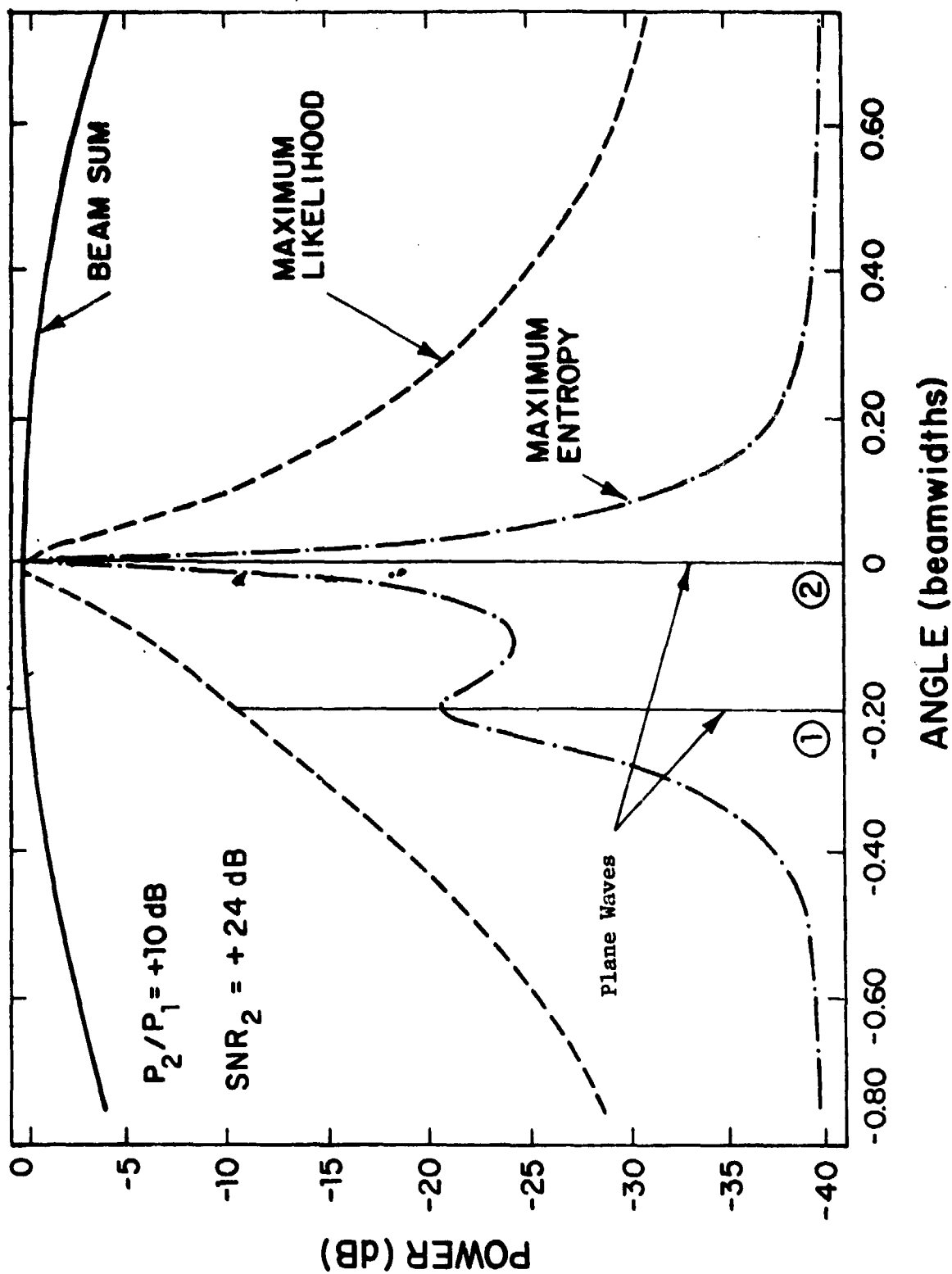
Fig. 5-13. Example of resolution of non-equal power incoherent plane waves.

102206-R

5-28
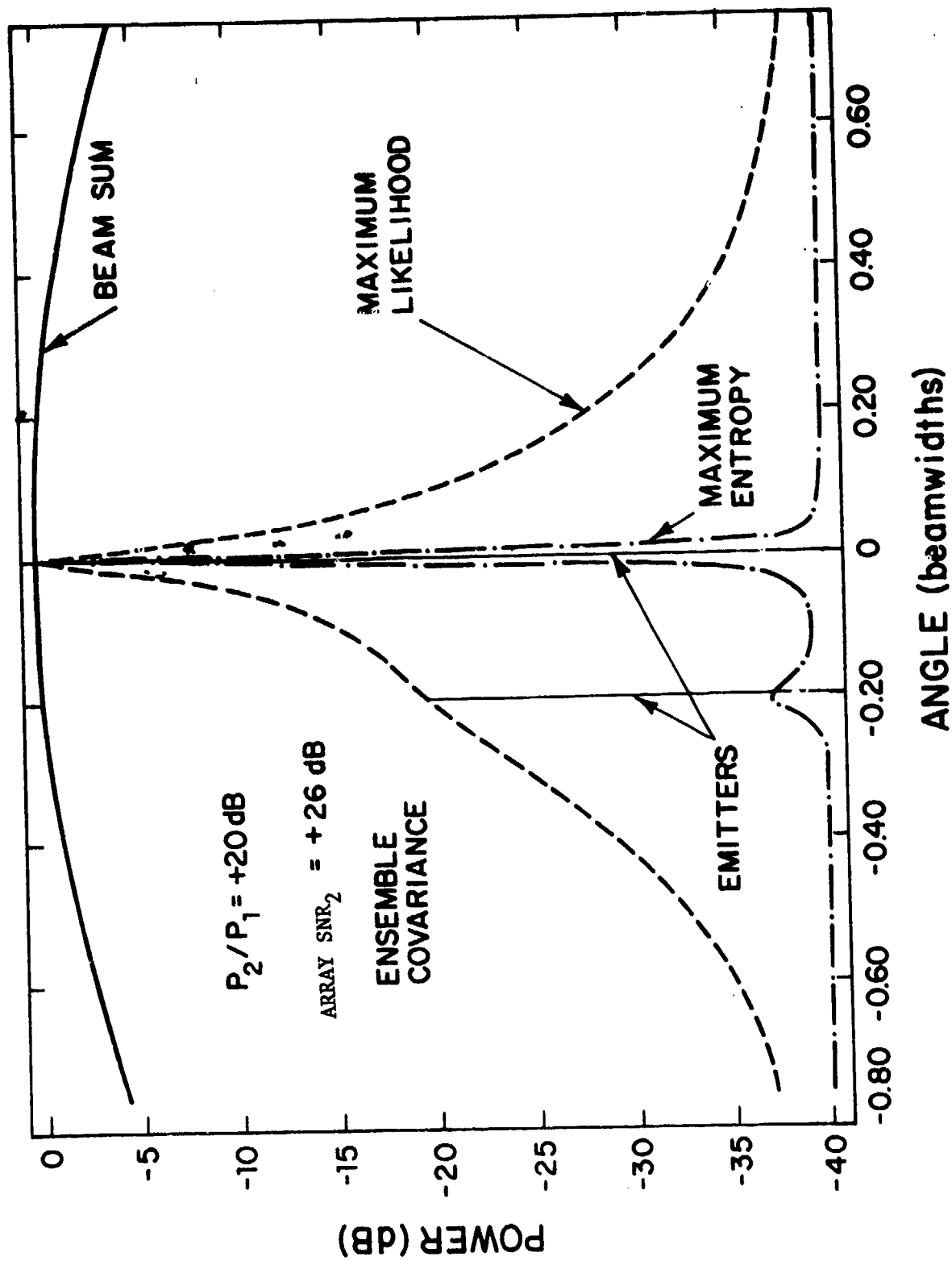
Fig. 5-14. Example of resolution of non-equal power incoherent plane waves.
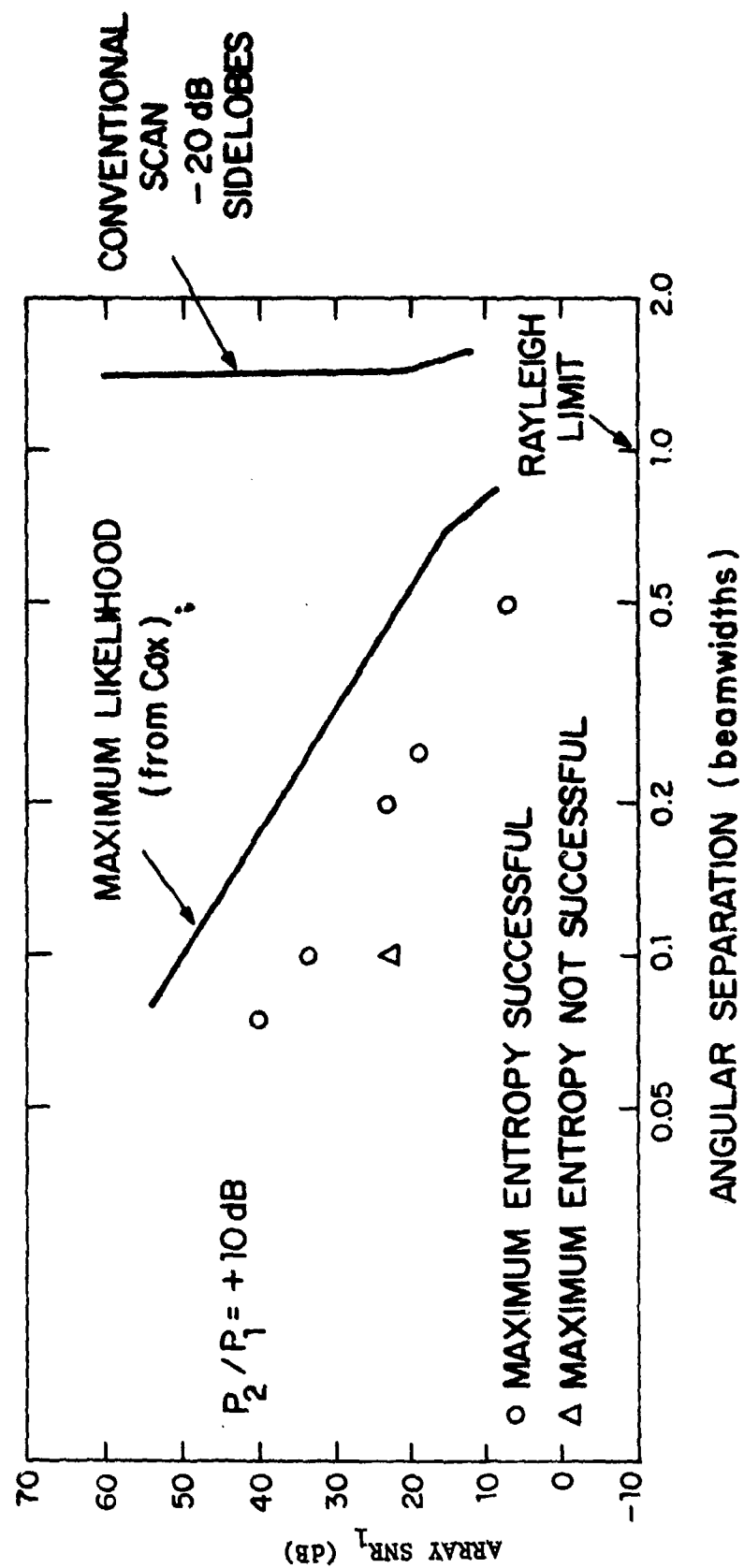
Fig. 5-15. Comparison of angular spectrum resolution limits for two incoherent plane waves with 10dB power difference.
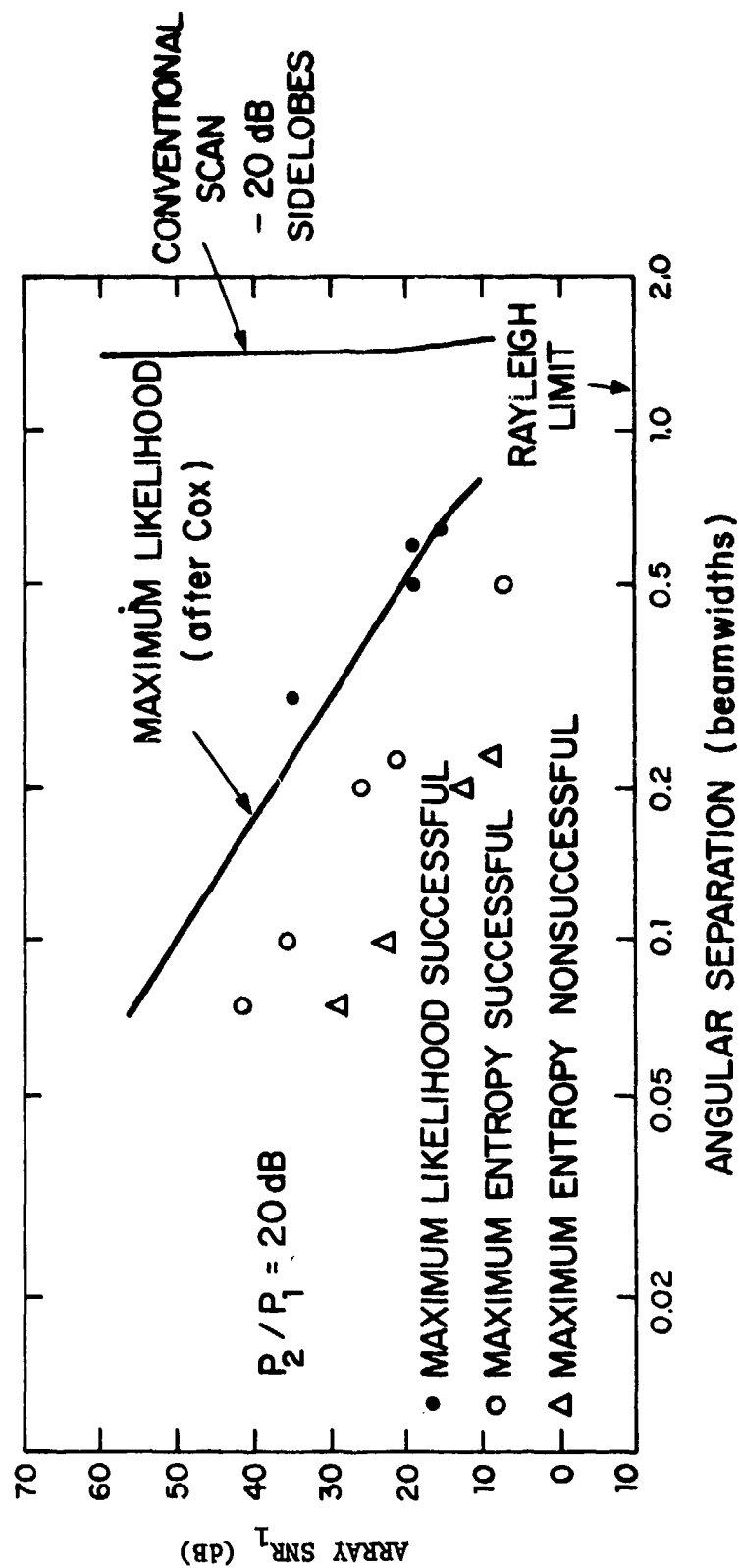
Fig. 5-16. Comparison of angular spectrum resolution limits for two incoherent plane waves with 20dB power difference.
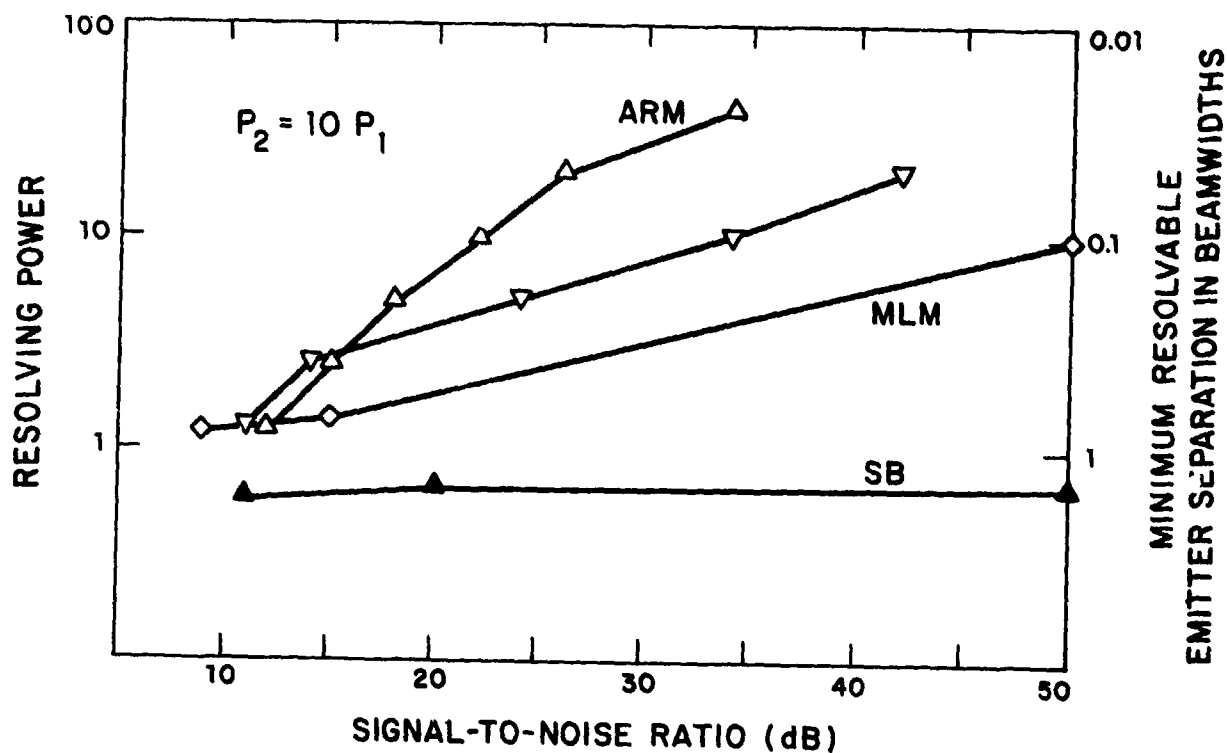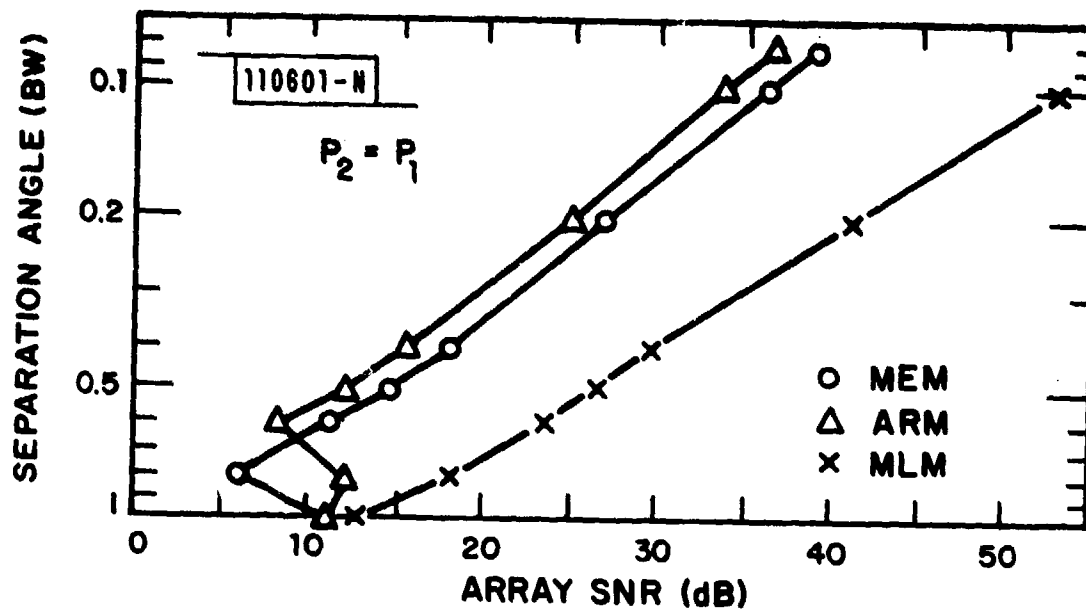
Fig. 5-17. Comparison of resolution limits for incoherent plane waves.

5-32

6 Elements
0.2 BW Separation
$P_2/P_1 = 10$ dB

X = infinite SNR
root location

O = finite SNR
root location

SNR=20 dB

SNR=30 dB

SNR=15 dB

SNR=25 dB
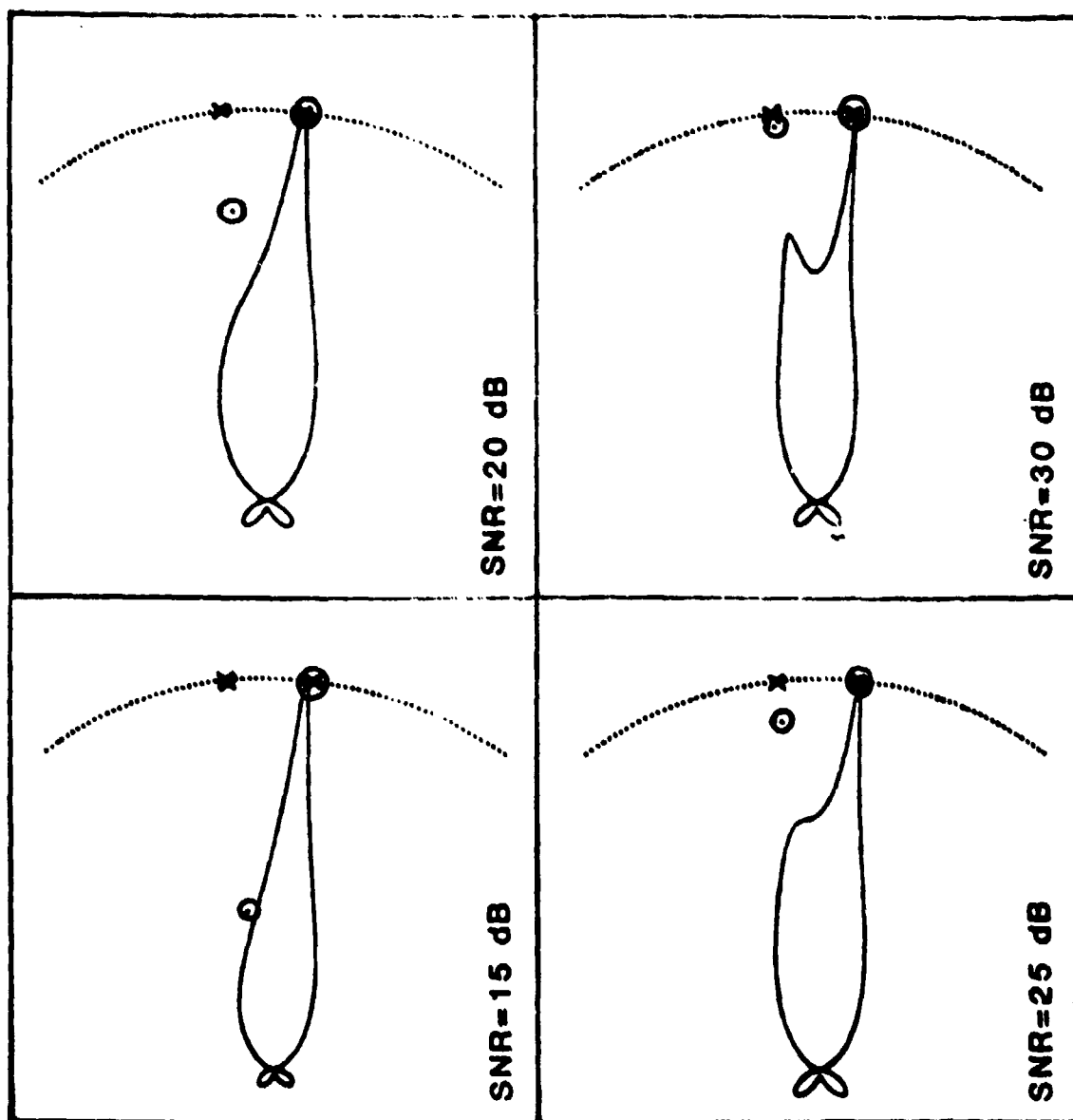
Fig. 5-18. ME/AR error filter root locations and corresponding spectral estimates for incoherent plane waves.
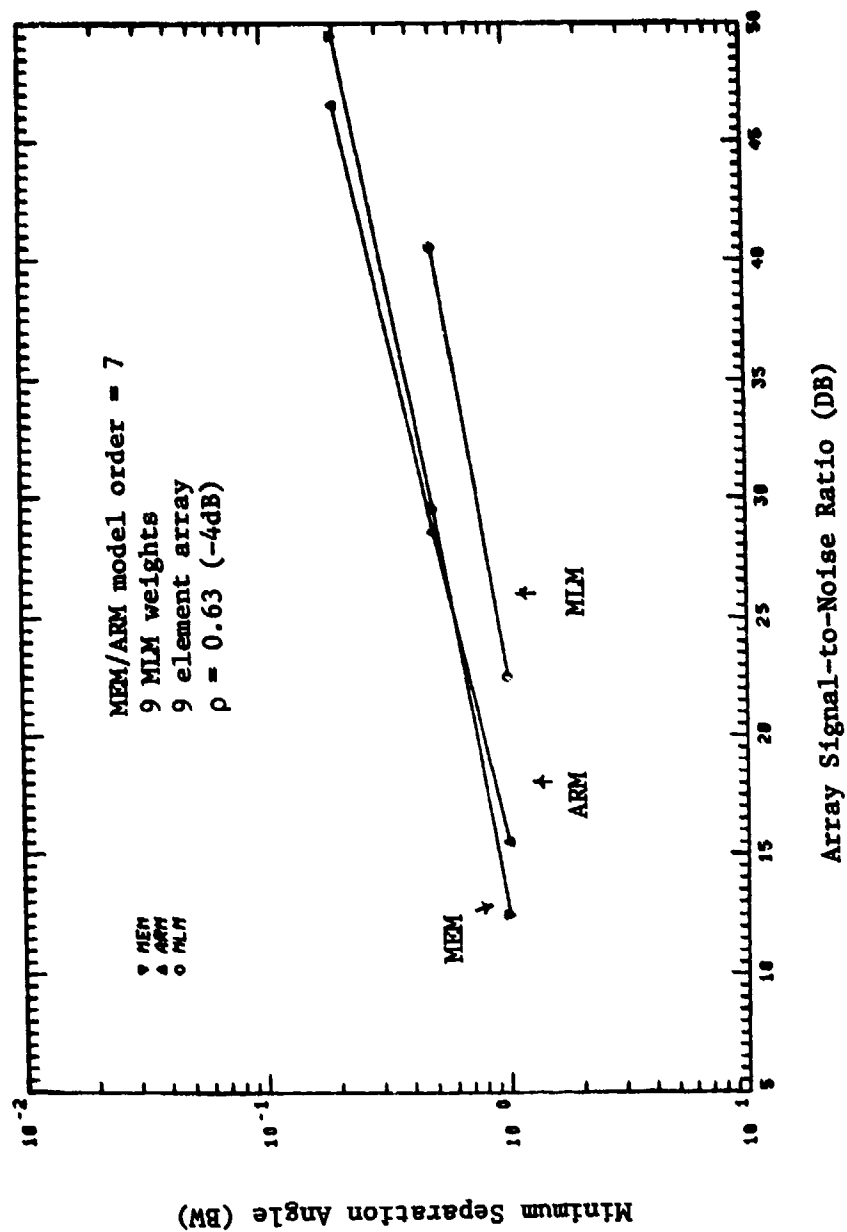
Fig. 5-19. Resolution threshold for two inphase coherent plane waves.

RELATIVE PHASE = 90 DEGREES

Minimum Separation Angle (BW)

Array Signal-To-Noise Ratio (DB)

MEM/ARM model order = 7
9 MLM weights
9 element array
ρ = 0.63 (-4dB)

MEM

MLM

ARM
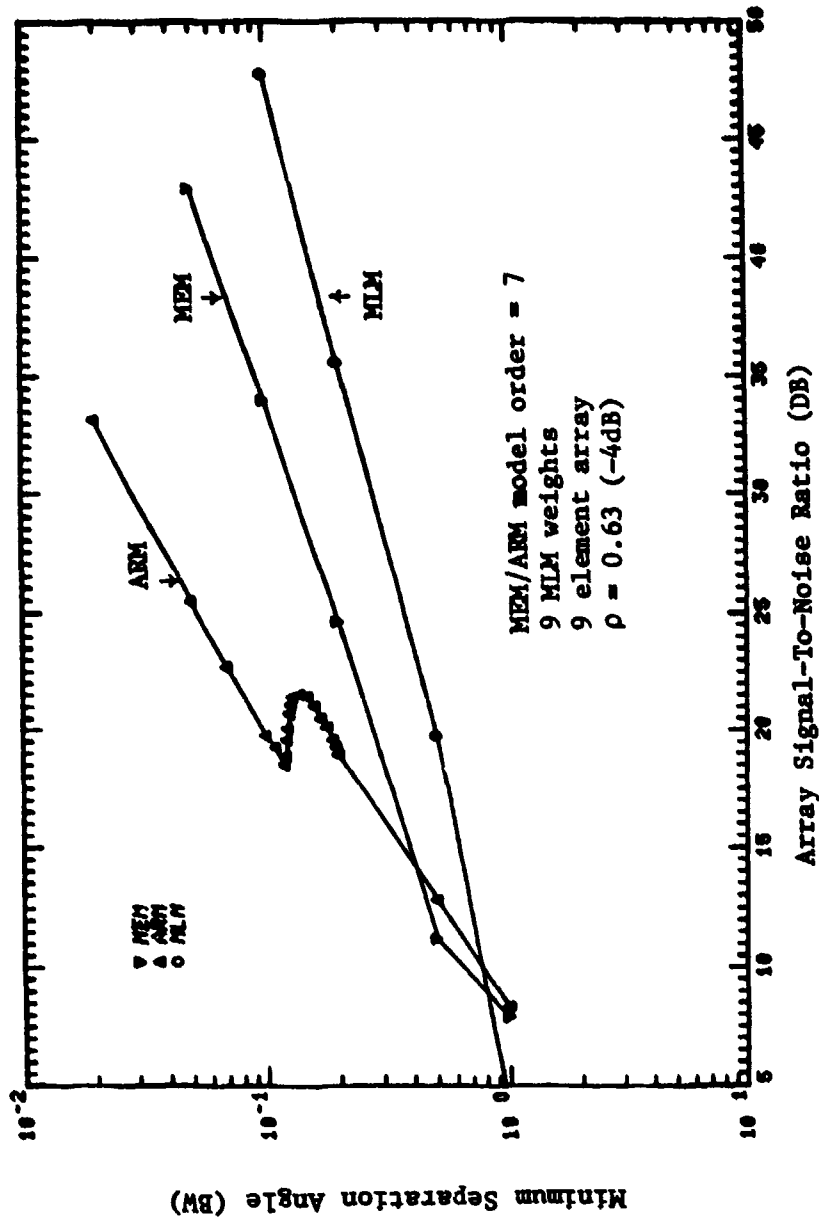
Fig. 5-20. Resolution threshold for two quadrature phase coherent plane waves.

RELATIVE PHASE = 180 DEGREES

MEM/ARM model order = 7
9 MLM weights
9 element array
ρ = 0.63 (-4dB)

Minimum Separation Angle (BW)

Array Signal-To-Noise Ratio (DB)
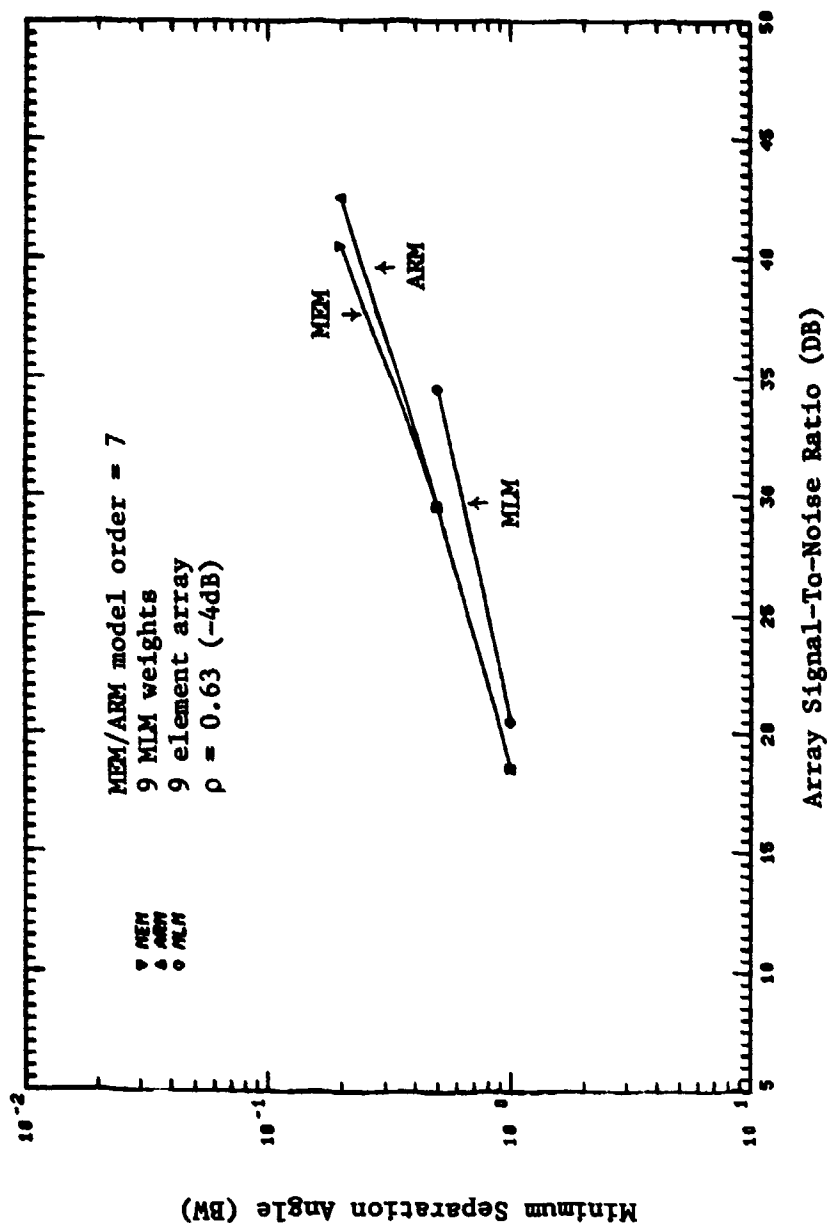
Fig. 5-21.   Resolution threshold for two out of phase coherent signals.

5-36
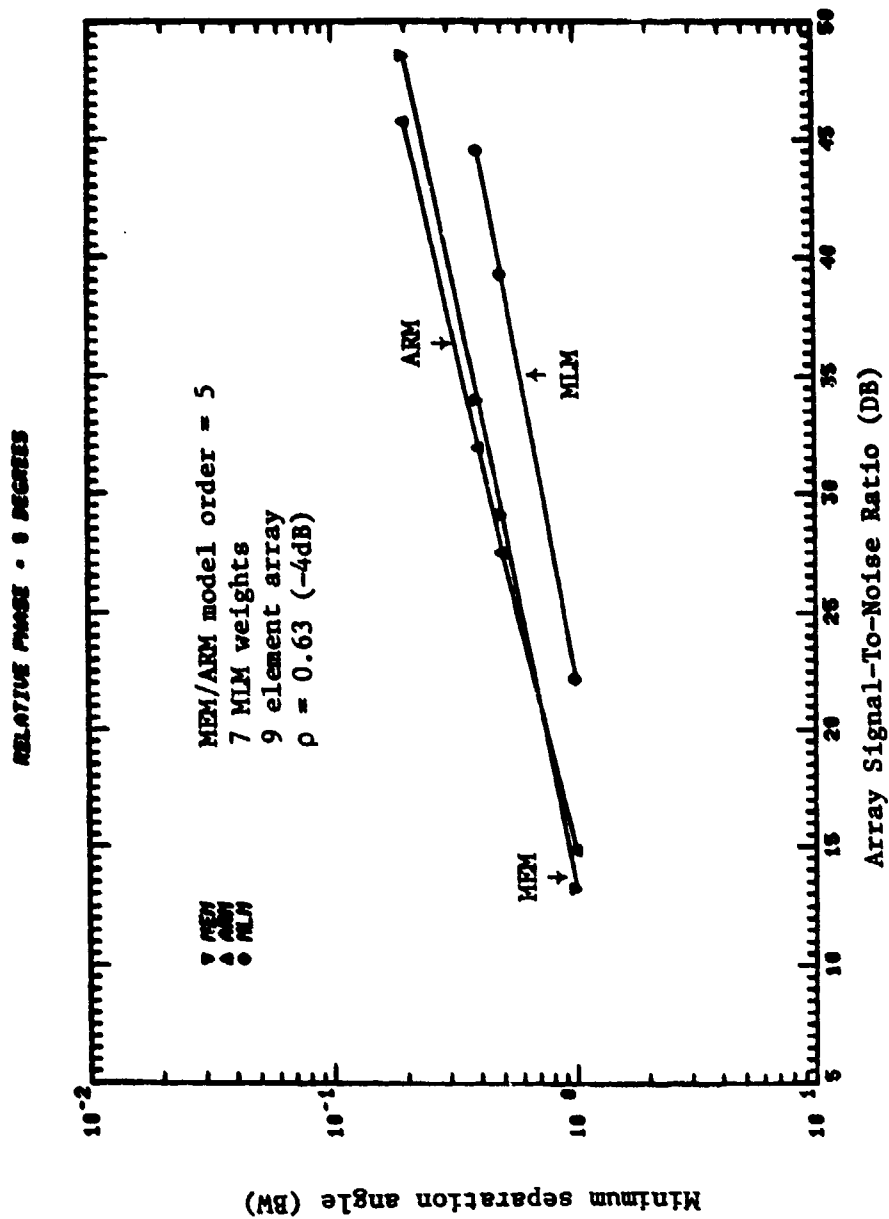
Fig. 5-22. Resolution threshold for two inphase coherent plane waves.
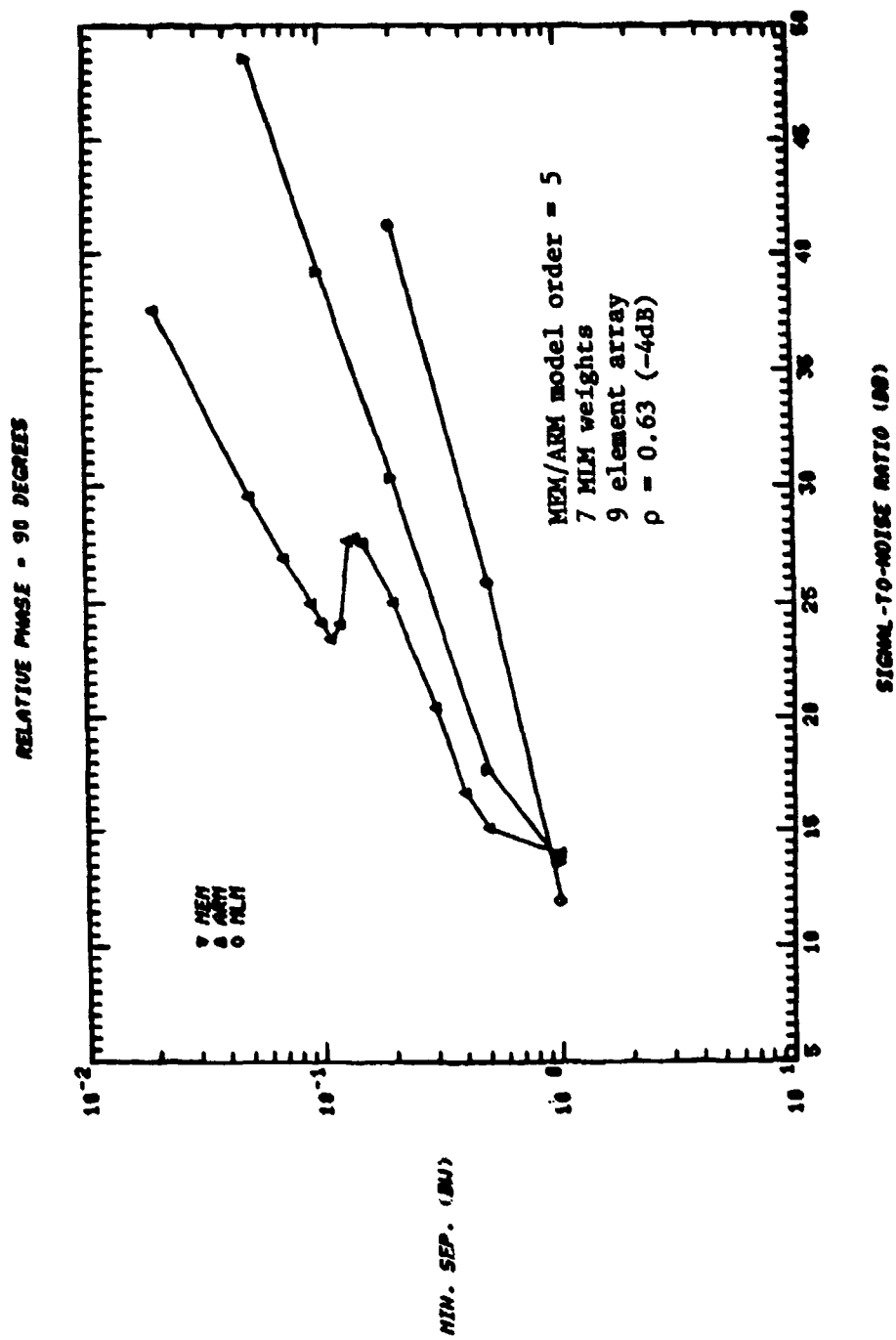
Fig. 5-23. Resolution threshold for two quadrature phase coherent plane waves.
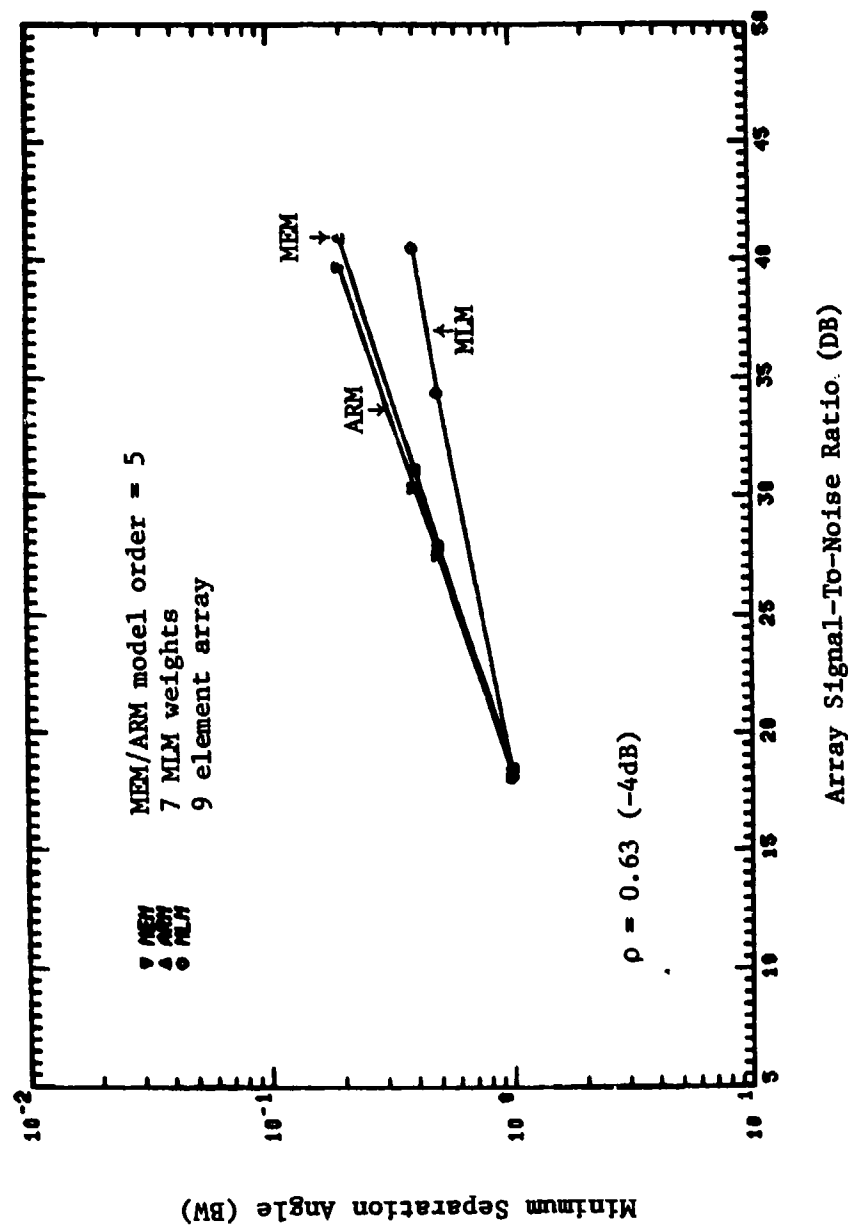
Fig. 5-24. Resolution threshold for two out of phase coherent plane waves.

results for the same signal environment, but lower model orders. Forward-backward smoothing was used in all cases. The lower model order (i.e., greater degree of spatial averaging) is seen to be beneficial in both cases.

The resolution criteria used here were similar to those used for the incoherent case:

(a) a proper number of peaks (or root locations) which are at least 3 dB above the nominal noise level,

(b) angular error < the lesser of (1/30 BW, 1/3 angular separation).

The fractional angular separation criterion is primarily needed at small separation angles (i.e., < 0.1 BW) with a 90° phase difference.

The resolution thresholds for 0° and 180° relative phase are seen to be quite similar. However, the failure mechanism is quite different as shown in Figs. 5-25 and 5-26. The spectra for the inphase condition typically show a single peak located between the two plane-wave arrival angles. By contrast, separate peaks occur in the out-of-phase condition, but the peak locations are typically much further apart than the actual plane waves.

The rather odd "S" shaped threshold curve for a 90° phase condition is due to the angle error criteria. In some SNR regimes, the error changes sign as one decreases the separation angle at fixed SNR. This causes the resolution threshold to decrease in some cases as the separation angle decreases. By contrast, when the angle error near threshold is always of a given sign (as in the 0° or 180° cases), the resolution threshold increases monotonically as angular spectrum decreases.

2.   Three Plane-Wave Cases

With conventional Fourier techniques (using low sidelobe spectrum analysis windows), one can define a resolution cell which characterizes the minimum angular separation for resolving a number of plane waves. Since the various "high resolution" techniques utilize nonlinear operations on the received data, one cannot assume a priori that the resolution thresholds for three (or more) plane waves will be similar to those for two plane waves.

5-40
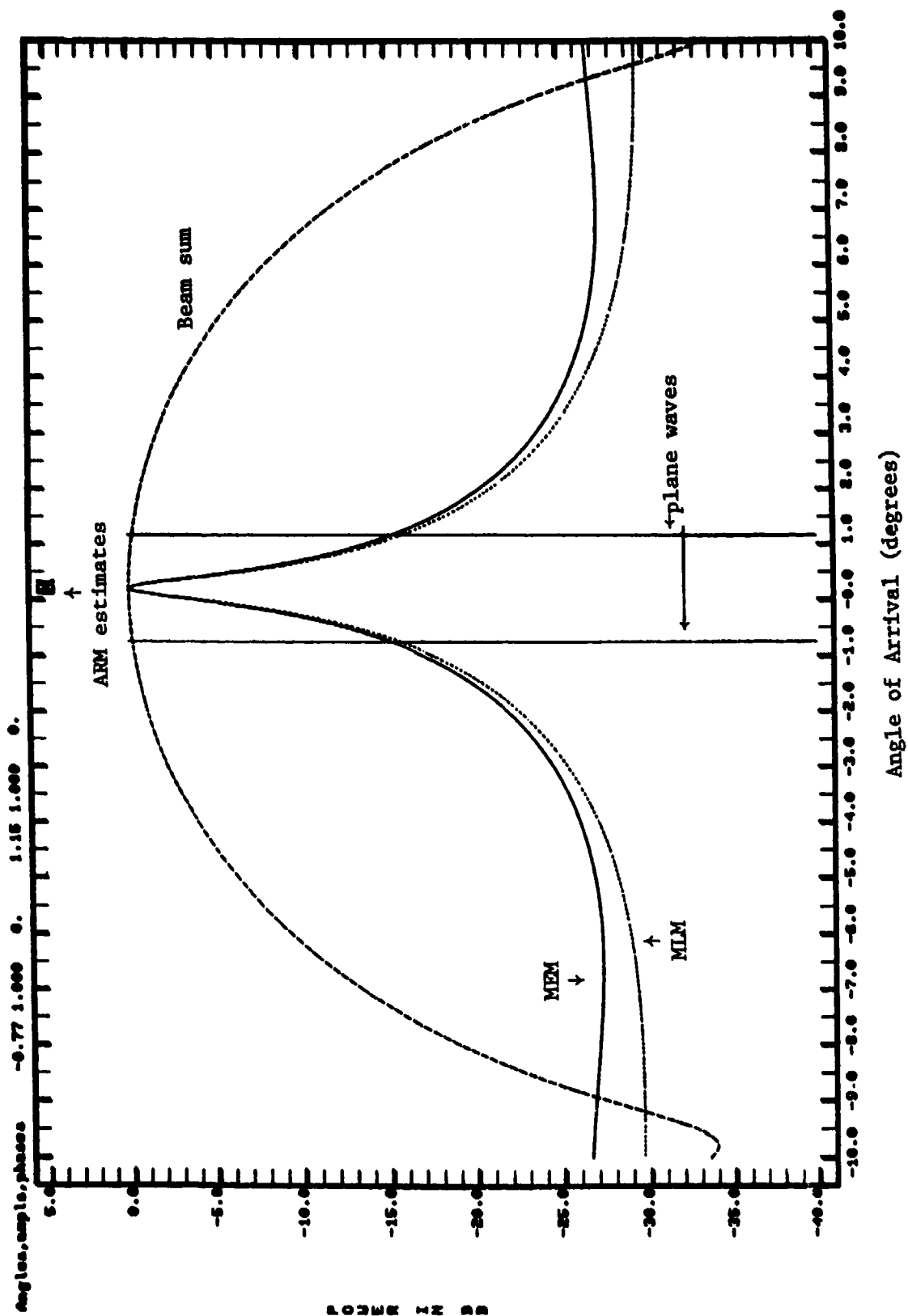
Fig. 5-25. Example of spectra for unresolved inphase coherent plane waves.

5-41
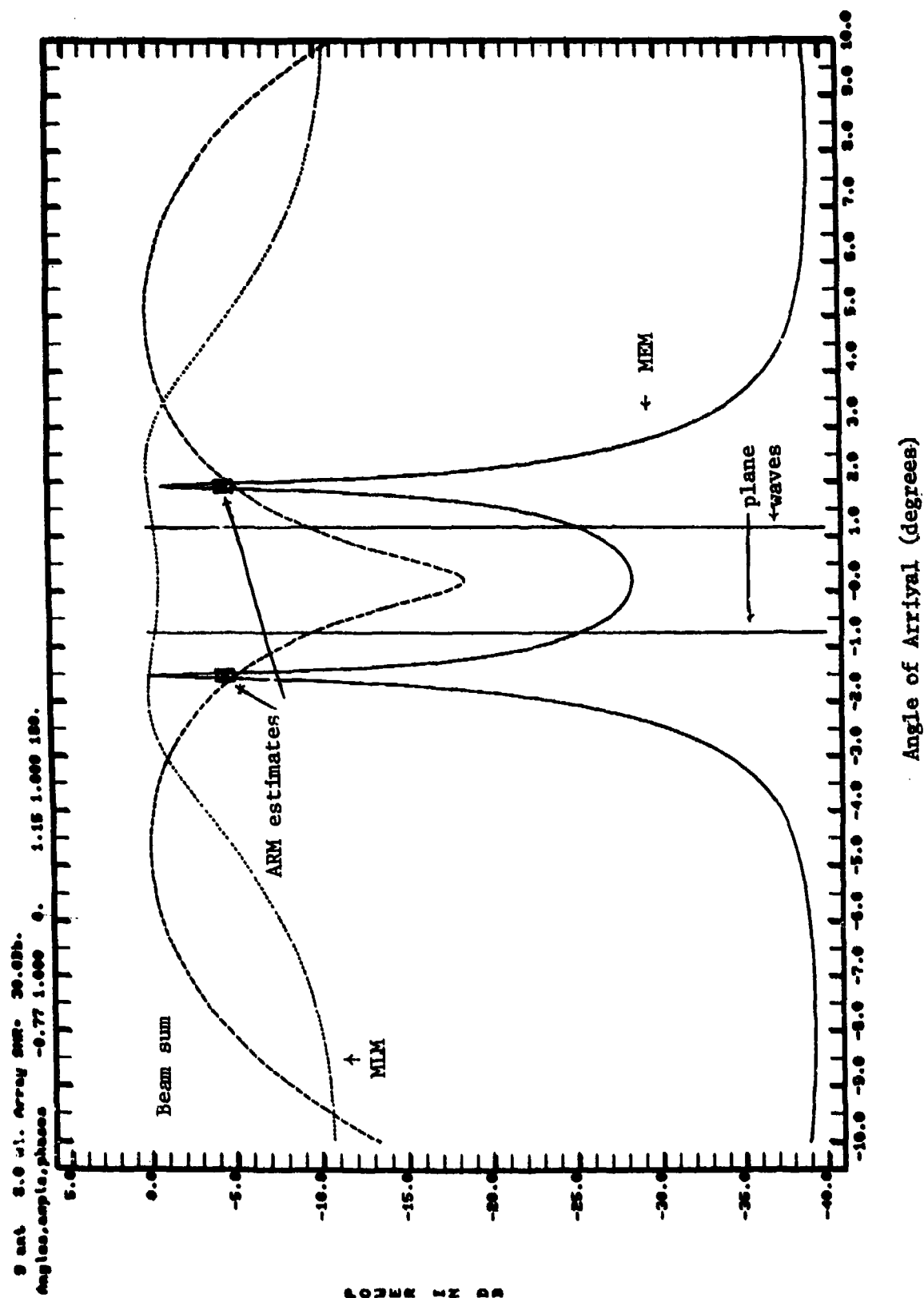
Fig. 5-26. Example of spectra for unresolved out-of-phase coherent plane waves.

Thus, a limited number of simulations were conducted to determine the extent to which three signals could be successfully resolved using representative high-resolution techniques.

The large number of possible amplitude/arrival angle/relative phase relationships for three plane waves made it impossible to accomplish a complete assessment. Rather, we have attempted to investigate two "canonical" cases and two cases which relate to typical MLS/ATC surveillance multipath problems. Figure 5-27 shows the plane wave amplitude and angle-of-arrival relationships for the four cases considered. In all cases, the separation angle parameter θ was the independent variable with the array SNR required to reach threshold acting as the dependent variable. For the two canonical cases, the threshold criteria were identical to those used in the two plane wave cases; namely an angular error less than the minimum of (1/30 BW, 1/3 θ). For the two multipath-related problems, the looser criteria of 0.1 BW angular error was used on the grounds that the current system performance in such cases is so poor that 0.1 BW errors would represent a substantial improvement.

Figures 5-28 and 5-29 compare the uncorrelated and correlated inphase signal performance for canonical case 1. In both cases, the presence of a third signal causes the resolution thresholds to rise substantially over that for two signals with the same separation. For both the correlated and uncorrelated signals, thresholds for three equally spaced plane waves are typically about twice those for two plane waves with the same spacing. The MLM performance degradation is approximately the same in dB as that for MEM at the same angular separation.

Figure 5-30 compares the performance with correlated and uncorrelated signals for canonical case 2. Here, we are interested in determining whether a third plane wave, which is quite resolvable by classical criteria, will unduly effect the ability to resolve two closely spaced plane waves. Comparing Fig. 5-30 with the corresponding two plane wave results (Figs. 5-12 and 5-19), we see that the existence of the third plane wave had relatively little effect on the MEM and MLM thresholds, but does cause some degradation in ARM performance against uncorrelated signals.

AMPLITUDE

|←—θ—→|←—θ—→|

ARRIVAL ANGLE (BW)

(a) CANONICAL CASE 1: EQUALLY SPACED, EQUAL AMPLITUDE PLANE
WAVES

AMPLITUDE

|←—θ—→|←——— 4 BW ———→|

ARRIVAL ANGLE (BW)
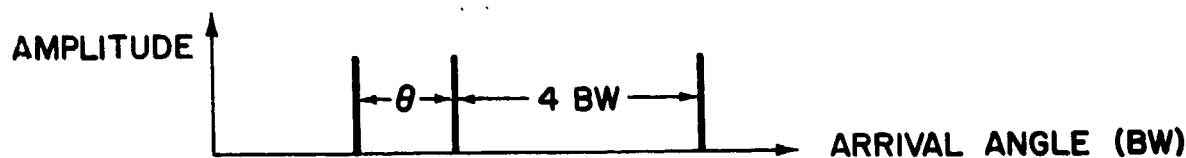
(b) CANONICAL CASE 2: "RESOLVABLE" PLANE WAVE WITH TWO CLOSELY
SPACED PLANE WAVES

AMPLITUDE

|←—θ—→|←—θ—→|

O

ARRIVAL ANGLE (BW)
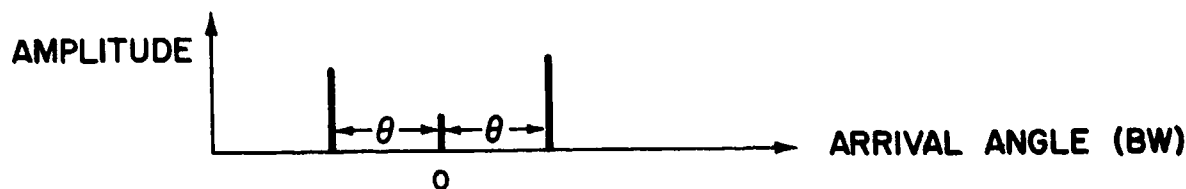
(c) MULTIPATH CASE 1: REFLECTIONS NEAR HORIZON AND IMAGE ANGLE

AMPLITUDE

|←θ→|←θ→|

ARRIVAL ANGLE (BW)
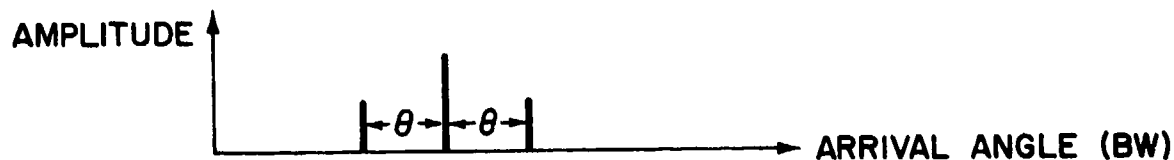
(d) MULTIPATH CASE 2: DIFFRACTION SIGNALS ON EITHER SIDE OF DIRECT
SIGNAL

Fig. 5-27. Three plane wave cases.

ARRAY SNR (dB)

O MEM
X MLM
△ ARM

$P_1 = P_2 = P_3 = 1$
$\theta_3 = -\theta_1 = \theta$
$\theta_2 = 0$

θ (BW)

0.1
0.2
0.5
1.0
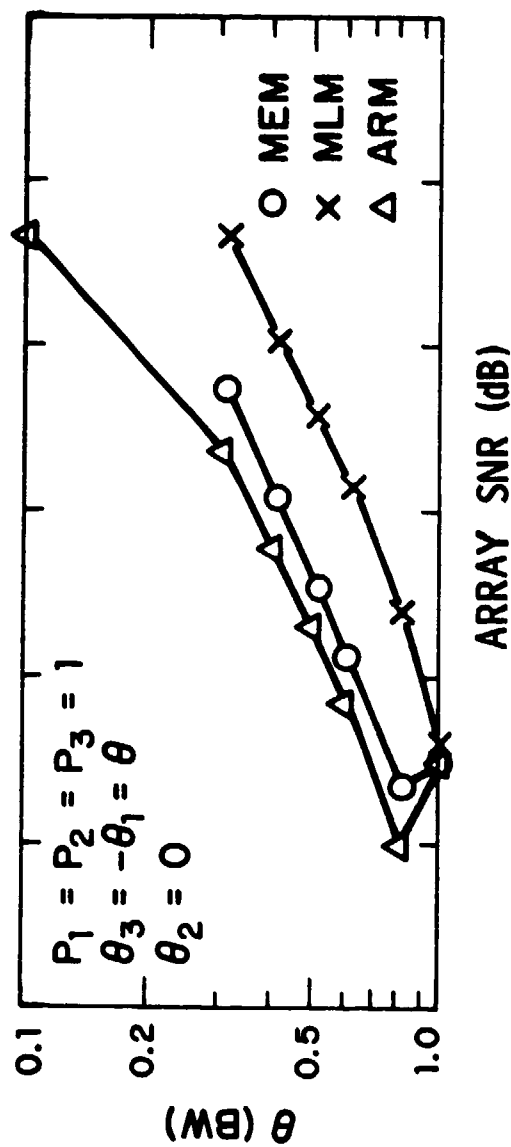
Fig. 5-28. Resolution threshold for three incoherent signals – canonical Case 1.

ARRAY SNR (dB)

0   10   20   30   40   50   60

$P_1 = P_2 = P_3 = 1$
$\theta_3 = -\theta_1 = \theta$ $\theta_2 = 0$
$\phi_1 = \phi_2 = \phi_3 = 0$

θ (BW)

0.1
0.2
0.5
1.0
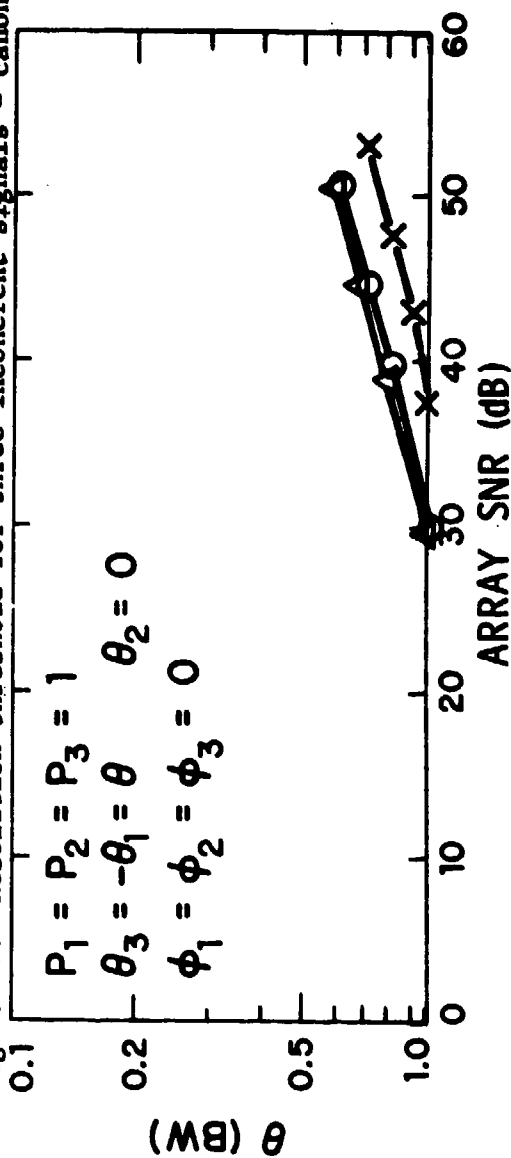
COHERENT INPHASE SIGNALS

Fig. 5-29. Resolution threshold for three signals Canonical Case 1.

UNCORRELATED SIGNALS

$P_1 = P_2 = P_3 = 1$
$\theta_1 = 0 \quad \theta_2 = +\theta$
$\theta_3 = \theta + 4BW$

O MEM
X MLM
△ ARM

COHERENT SIGNALS

$P_1 = P_2 = P_3 = 1$
$\theta_1 = 0 \quad \theta_2 = +\theta \quad \theta_3 = \theta + 4BW$
$\phi_1 = \phi_2 = \phi_3 = 0$

ARRAY SNR (dB)

θ (BW)
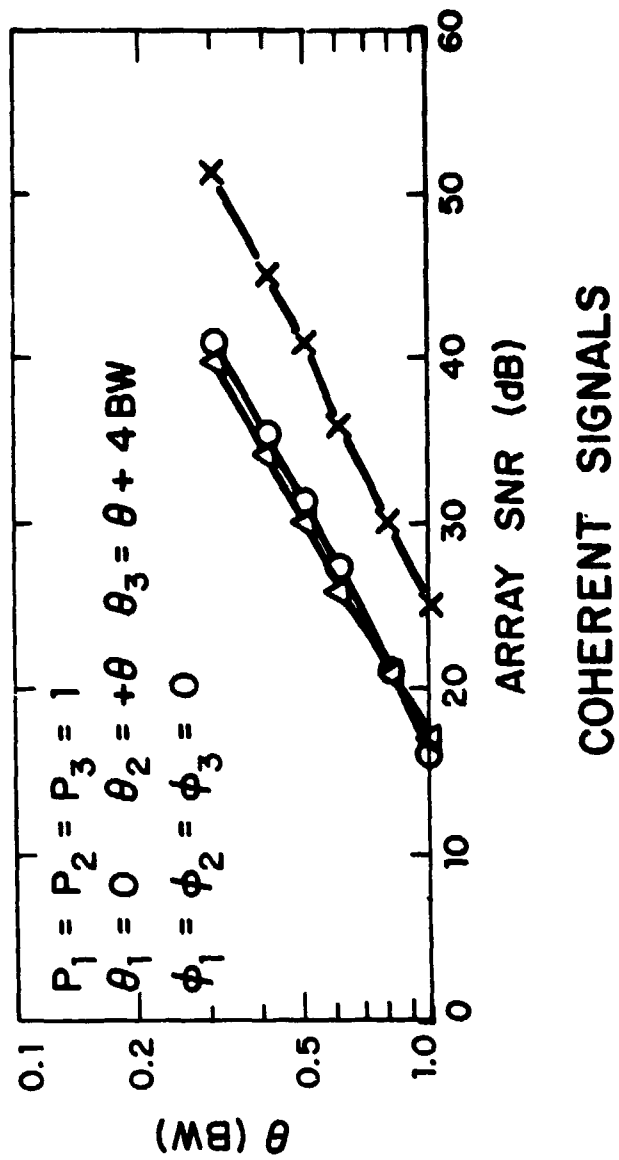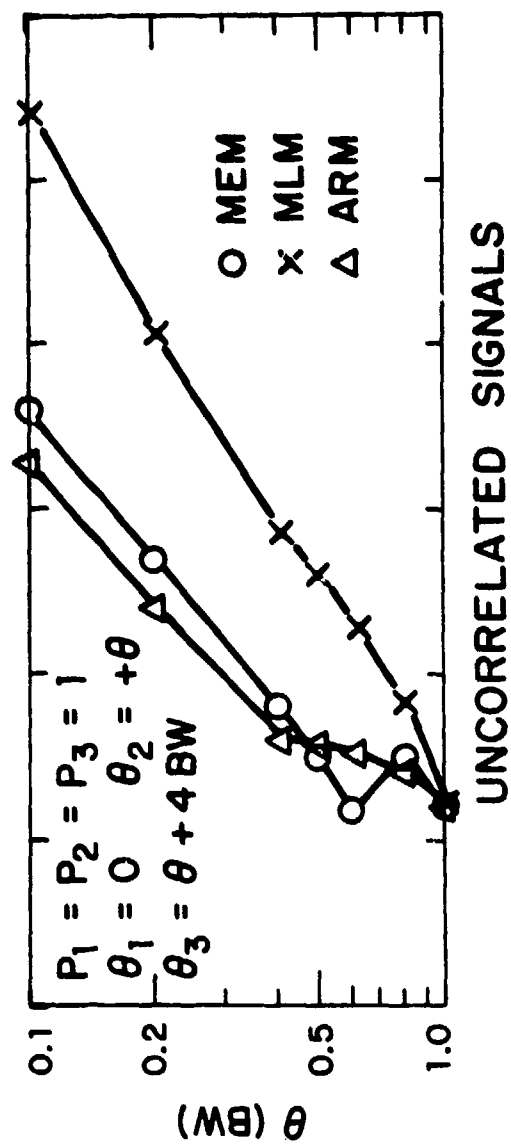
Fig. 5-30.   Resolution threshold for three signals canonical Case 2.

Figures 5-31 and 5-32 compare the correlated and uncorrelated signal resolution performance for multipath case No. 1. This case was intended to correspond to the occurrence of elevation multipath consisting of specular reflections at the image angle and from near the horizon as well as the direct signal at a positive elevation angle. The plane wave near the horizon could be viewed as a (very) crude approximation to diffuse multipath from an extended region near the horizon [14]. The principal problem here is identification of the center plane wave, since there are larger plane waves to either side.

At 1 BW separation, the uncorrelated resolution thresholds are fairly similar to those for two-signal resolution. However, at 0.5 BW separation angle, the SNR required for MLM and MEM are some 10 dB higher for the three-signal case, but ARM is similar. At smaller separation angles, the differences are quite large. Straight-line fits to the curves yielded the following slopes:

| | dB/decade of separation angle | |
	two plane waves	three plane waves
MEM	15	50
ARM	28	52
MLM	40	58

The two correlated signal cases show roughly similar slopes; however, the case where the middle plane wave is in quadrature with the other two plane waves has a threshold which is 5 to 12 dB lower than the "all in-phase" case. It would appear from Fig. 5-32 that for typical field measurement SNR values of 40 dB, it will be difficult to resolve elevation multipath signals near the horizon for direct signal angles of less than 0.7 BW.

Figures 5-33 and 5-34 compare the correlated and uncorrelated resolution performance for multipath case No. 2. This case is intended to roughly represent a common azimuth shadowing problem wherein the direct signal diffracts over the top of the obstruction while interfering signals diffract around the side of the obstacle. The uncorrelated and coherent signal/in-phase thresholds are quite similar to canonical case No. 1, as would be expected, since
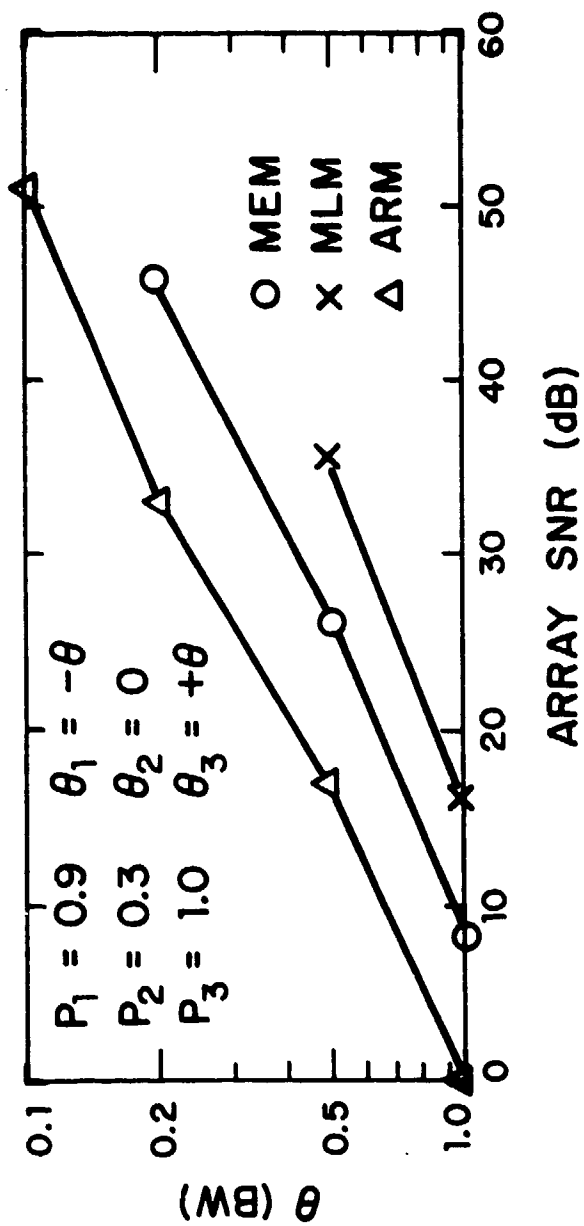
Fig. 5-31. Three uncorrelated signal resolution threshold multipath Case No. 1.
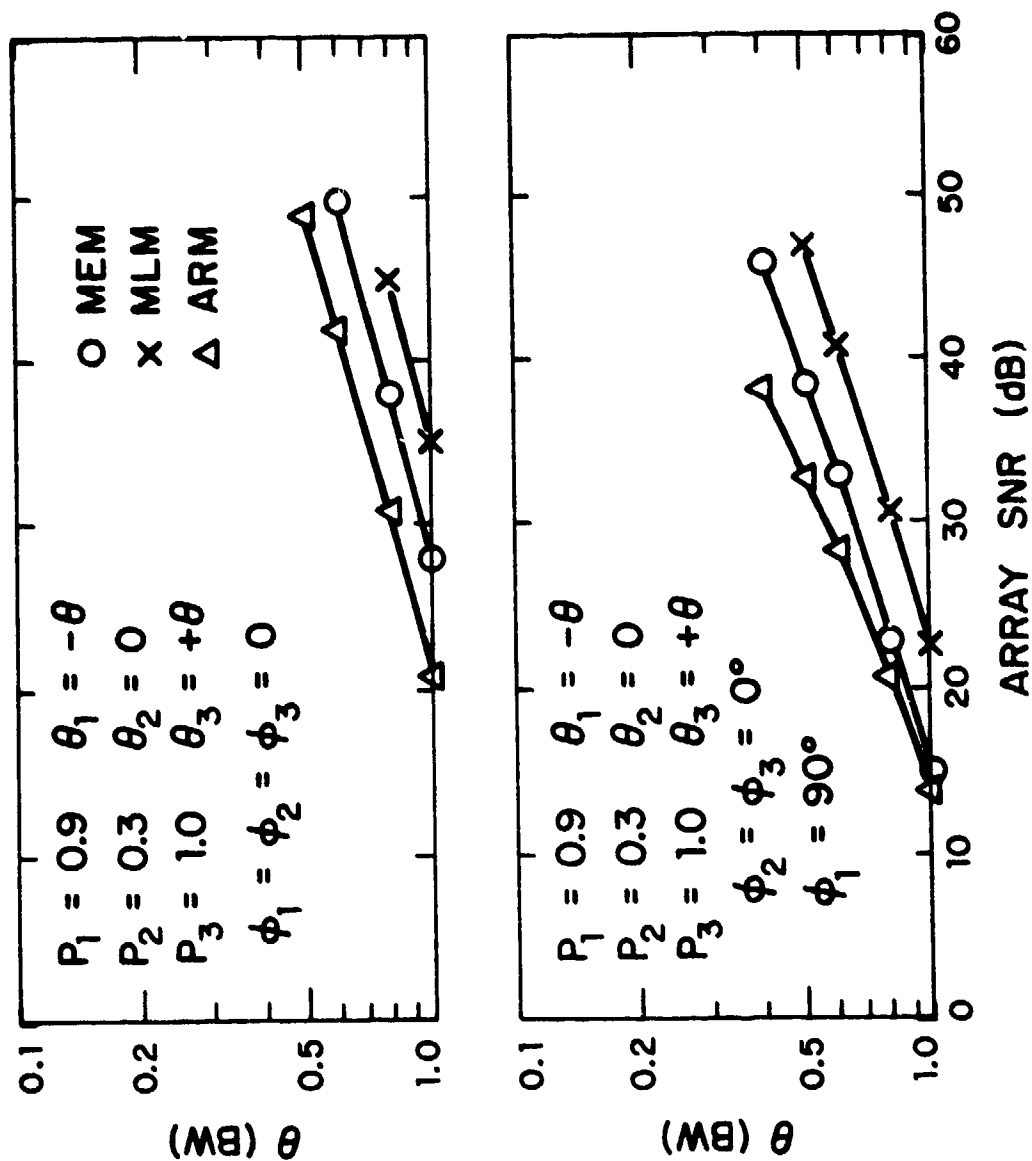
Fig. 5-32. Resolution threshold for three correlated signals multipath Case 1.
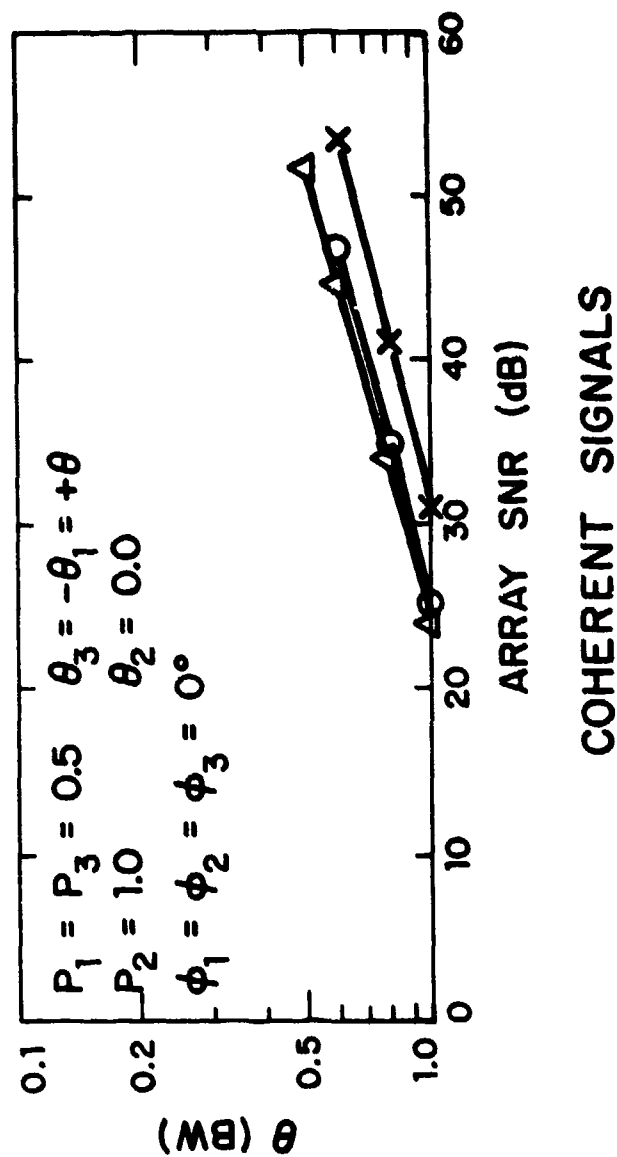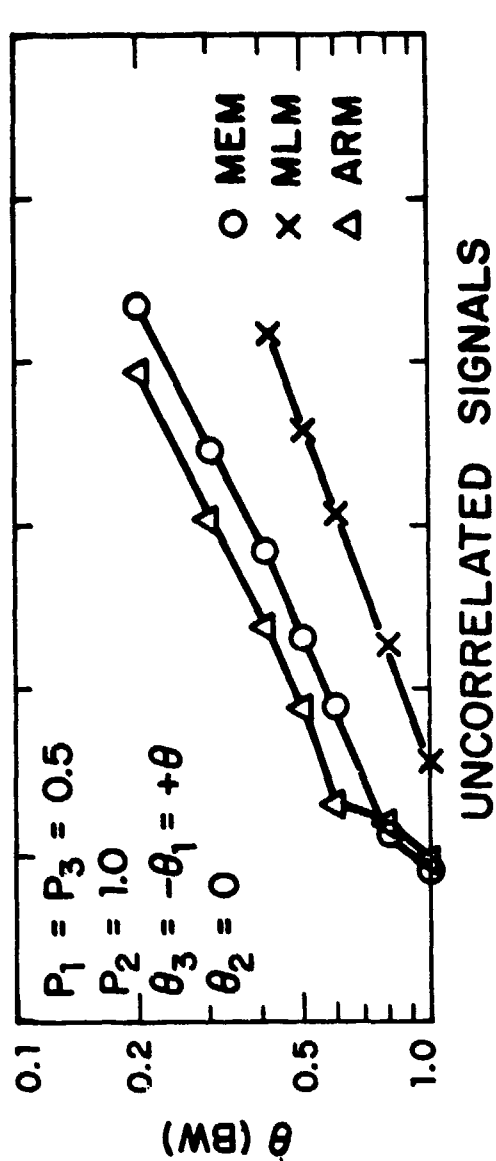
Fig. 5-33. Resolution threshold for three signals multipath Case 2.
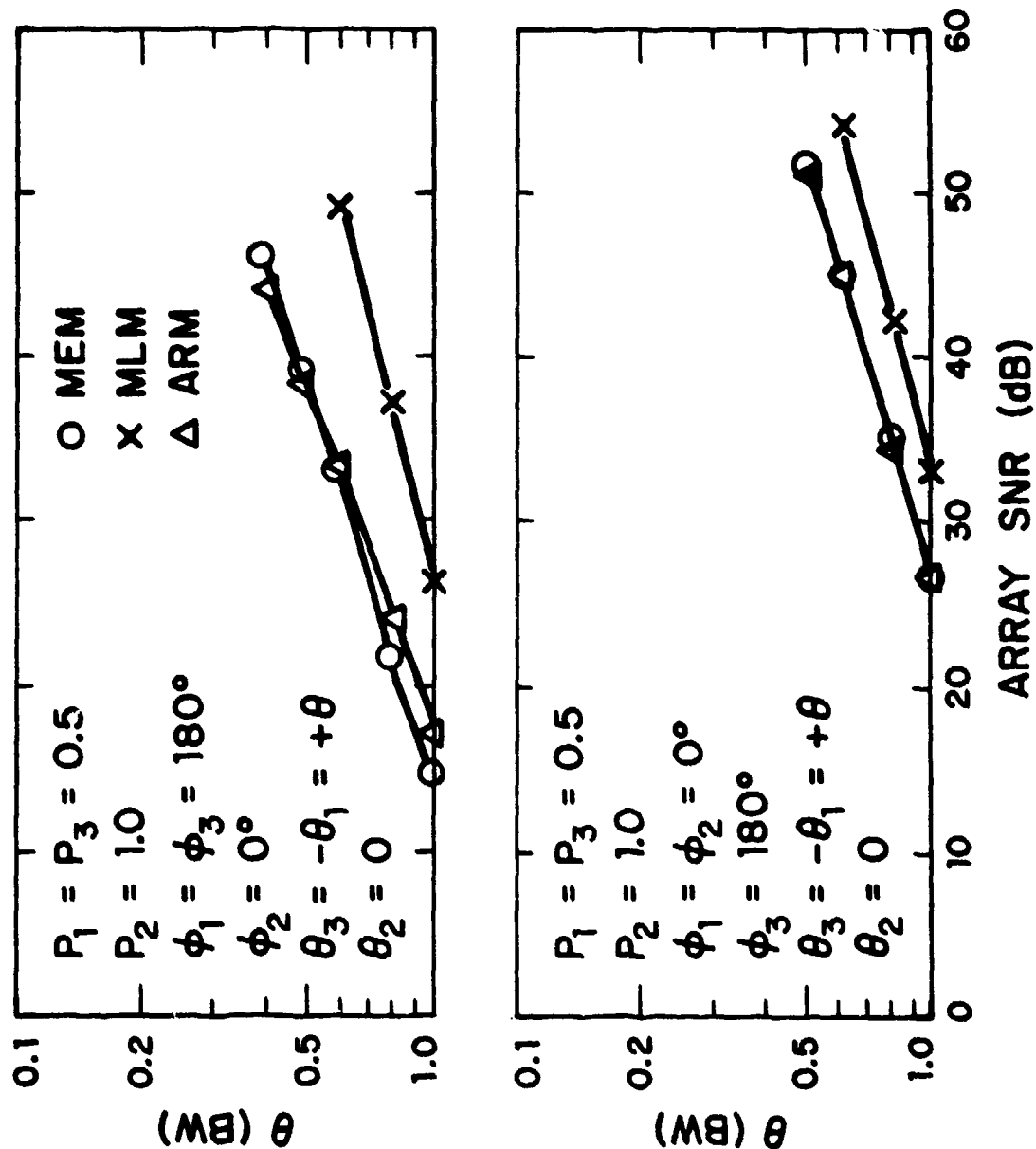
Fig. 5-34. Resolution threshold for three coherent signals multipath Case 2.

the only difference is in the amplitudes of the two outer plane waves. The most difficult case for conventional system angle-of-arrival estimation is when one or both of the outer signals are out of phase with the center signal. Comparing Figs. 5-33 and 5-34, we see that the out-of-phase conditions yield resolution thresholds which are similar to or lower than the inphase case.

D. Experimental Data Comparison

Four processing techniques (MEM, MLM, MUSIC, and ARM) discussed in the previous chapter were applied to two groups of field data. Representative results are presented here. The first group of field data was obtained from a series of experiments at the Lincoln Laboratory Antenna Test Range, where our purpose was to resolve two azimuthally distributed signal sources. Here, the signal environment was completely known, since we set up two signal sources to radiate with known power ratio, signal-to-noise ratio (SNR), azimuthal angle separation, and relative phase. The second group of field data was obtained from various field measurements at several test sites of different terrain types where our objective was to characterize terrain multipath environment. Here, the received signal environment was partially unknown. That is, we did not know the number of multipath signals, their angles of arrivals, and their relative signal power.

As noted in the previous section, there are several computational options available to a given processing technique, and in general, different options will yield different results. For example, different MLM results can be obtained with various options of estimating the covariance matrix. The results presented below were obtained with those options which we judged to be suitable for a given estimation technique when only a single snapshot was available. For MEM and ARM, the prediction error filter coefficients were computed based on the modified covariance method with the filter length (LRA) set around $L/3+1$ (L is the number of sensor elements). For MLM, the modified covariance matrix was used with the order of around $L/2+1$. For MUSIC, the unbiased covariance estimate was used in the first group of field data and the modified covariance estimate with the order of around $2L/3$ was used in the

5-52

second group of field data. The number of eigenvalues selected for MUSIC was the number of those eigenvalues which were greater than the apparent noise level. Table 5-3 summarizes the actual parameter values (i.e., LRA, etc.) used in various cases.

1. First Group of Field Data: Known Signal Environment

The experimental setup for obtaining this group of field data is shown in Fig. 5-35. The first signal source was set at $0°$ with respect to the receiving array boresight, and the second signal source was set at any desired angular separation (azimuth) with respect to the first signal source. Signals radiated from two sources were from the same signal generator, and the relative phase between two signals was varied from $0°$ to $360°$ during a given measurement. For aperture sampling of the incoming signal, we used a 19-element 56.7-wavelength (λ) aperture C-band array whose beamwidth (BW) was about $1°$. Here, we are interested in knowing (1) if various signal processing techniques can resolve these two signal sources at a given angular separation and (2) how the errors in the estimated azimuth angle of the signal sources vary with the relative RF phase between two sources.

Figures 5-36 and 5-37 show the results from the case #1 where angular separation (θ_{sep}) between two signal sources was 1 BW, the power ratio (P_1/P_2) was about 3 dB and the single antenna SNR was about 33 dB. Here, the estimated angle errors on the arrival angles of two signals were plotted as a function of the relative phase between two signal sources. Figure 5-36 is for the first signal source, and Fig. 5-37 is for the second signal source. For this case, all four techniques indicate the existence of two signal components in the incoming signal for all relative phases. MEM, MLM, and ARM yield similar angle errors for both signal sources. The peak angle error is around 0.07 BW for the first signal, and around 0.1 BW for the second signal. MUSIC shows much larger angle errors than the other techniques, especially for the second signal source. Also, the angle error from MUSIC shows much greater dependence on the relative phase between two signal sources. The larger angle errors occur around the relative phases of $\pm180°$ and $0°$ and the smaller errors occur around the relative phases of $\pm90°$.

5-53

TABLE 5-3

SUMMARY OF PARAMETER VALUES

	CASE	MEM, ARM LRA[1]	MLM NORD[2]	MUSIC NORD[3]	MUSIC #λ[4]
First group of data	Cases 1-4 (Fig. 5-2 to 5-9)		8	9	− 2
	Hanscom Airport L band (Fig. 5-11) C band (Fig. 5-12)		3 7	4 8	− 2 13 4
Second group of data	Fort Devens L band (Fig. 5-13) C band (Fig. 5-14)		5 8	5 9	6 3 13 4
	Otis AFB L band (Fig. 5-15) C band (Fig. 5-16)		5 8	5 9	6 3 13 4

(1) LRA: Prediction error filter length.
(2) NORD: Order of modified covariance matrix.
(3) For the first group of data and the L-band Hanscom Airport data, the unbiased covariance estimate was used. The modified covariance estimate was used for the rest of the cases.
(4) #λ: number of eigenvalues selected.
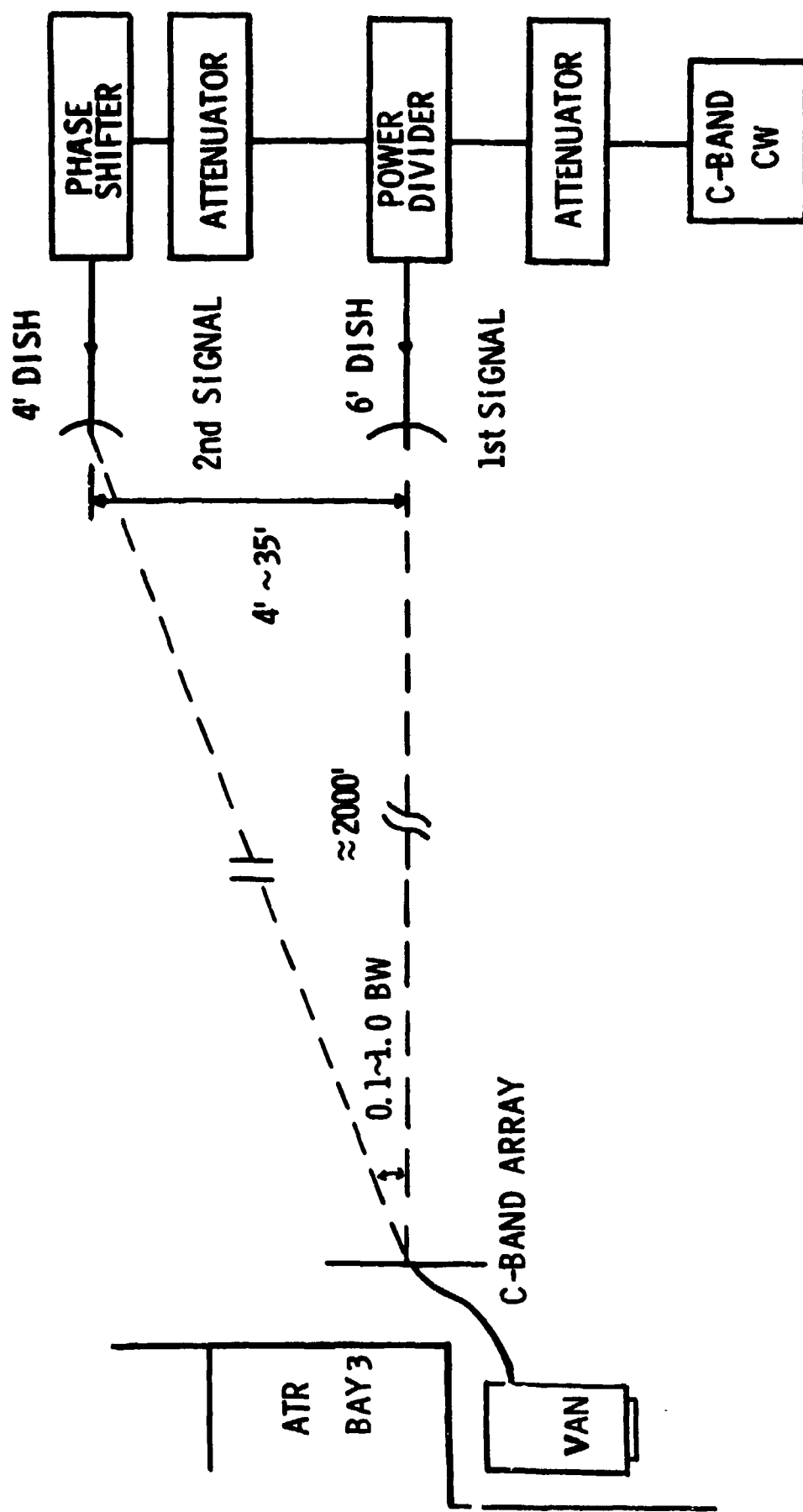
Fig. 5-35. Setup for two signal measurement.

First Signal

X = MEM
L = MLM
M = MUSIC
O = ARM

θ_{sep} = 1 BW
P_1/P_2 = 3 dB
SNR = 33 dB

RELATIVE RF PHASE ANGLE (DEG)
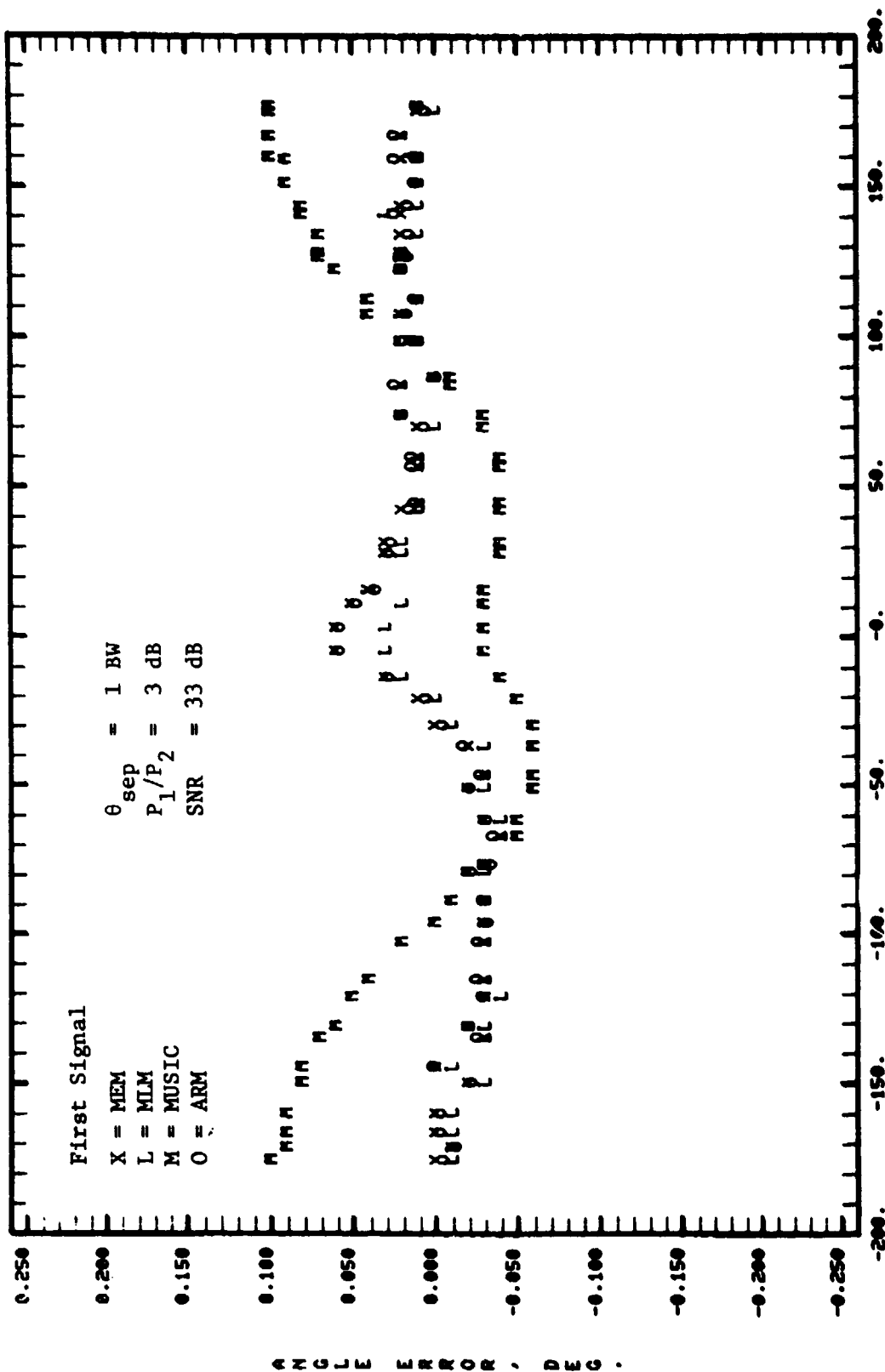
Fig. 5-36. Angle errors: first signal in Case 1.
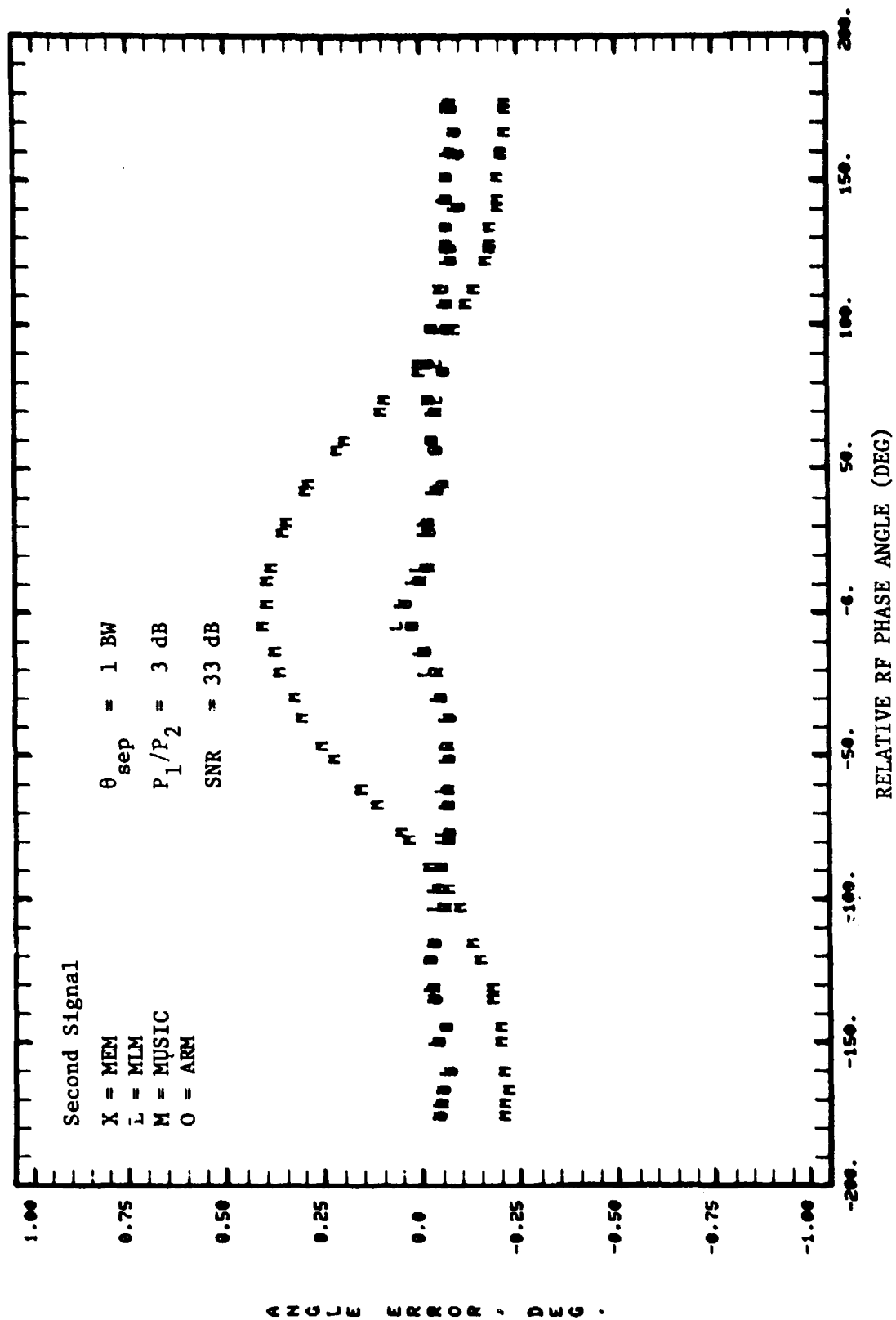
Fig. 5-37. Angle errors: second signal in Case 1.

Second Signal

X = MEM
L = MLM
M = MUSIC
O = ARM

θ_{sep} = 1 BW

P_1/P_2 = 3 dB

SNR = 33 dB

RELATIVE RF PHASE ANGLE (DEG)

ANGLE ERROR, DEG.

Figures 5-38 and 5-39 show similar results for case #2 where $\theta_{sep} = 1$ BW, $P_1/P_2 = 10$ dB and SNR ≈ 30 dB. Figure 5-38 gives the angle errors for the first signal, and Fig. 5-39 gives the corresponding errors for the second signal. Again, all four techniques can resolve two signal sources for all relative phases, although the second signal is much weaker in this case than in the previous case with the same θ_{sep} and SNR ($P_1/P_2 = 10$ dB versus 3 dB). For the first signal source, MEM, MLM, and ARM again result in similar angle errors, with the peak angle error of about 0.03 BW; while the MUSIC yields much larger errors (0.08 BW) for the relative phases near $\pm180°$. For the second signal source, MEM and ARM appear to give slightly smaller errors than MLM, 0.1 BW versus 0.15 BW; while MUSIC shows much larger errors of 0.25 BW near the $\pm180°$ relative phases and 0.55 BW around the $0°$ relative phase. As compared to the MEM, MLM, and ARM, the MUSIC angle error is much more sensitive to the relative phase between two signal sources.

Figures 5-40 and 5-41 show the results for the case #3, where $\theta_{sep} = 0.5$ BW, $P_1/P_2 \approx 3$ dB and SNR ≈ 35 dB. The angle errors for the first signal source are given in Fig. 5-40, and those for the second signal source are shown in Fig. 5-41. Here, only the ARM and MUSIC techniques correctly indicate the existence of two signal components in the incoming signal for all relative phases between two signals. The MEM technique only detects one signal for the relative phases around $0°$, and MLM does not yield any second signal at all. Contrary to the previous two cases, now the MUSIC angle errors are about the same as the ARM angle errors for both two signal sources. Also, now both the MUSIC and ARM angle errors show the strong dependence on the relative phase between two signals. For the first signal, both MUSIC and ARM show larger errors of 0.25 BW near $\pm180°$ relative phases and 0.15 BW around the $0°$ relative phase. For the second signal, MUSIC has larger errors of 0.45 BW near the $\pm180°$ relative phases, while ARM shows larger errors of 0.5 BW around the $0°$ relative phaes. For those relative phases where MEM can resolve the two signals, the MEM angle errors are about the same as those for ARM.

Figures 5-42 and 5-43 show the results for the case #4, where $\theta_{sep} = 0.5$ BW, $P_1/P_2 \approx 10$ dB and SNR ≈ 30 dB. Figure 5-42 gives the angle errors for the first signal, and Fig. 5-43 gives the corresponding errors for the second sig-

RELATIVE RF PHASE ANGLE (DEG)
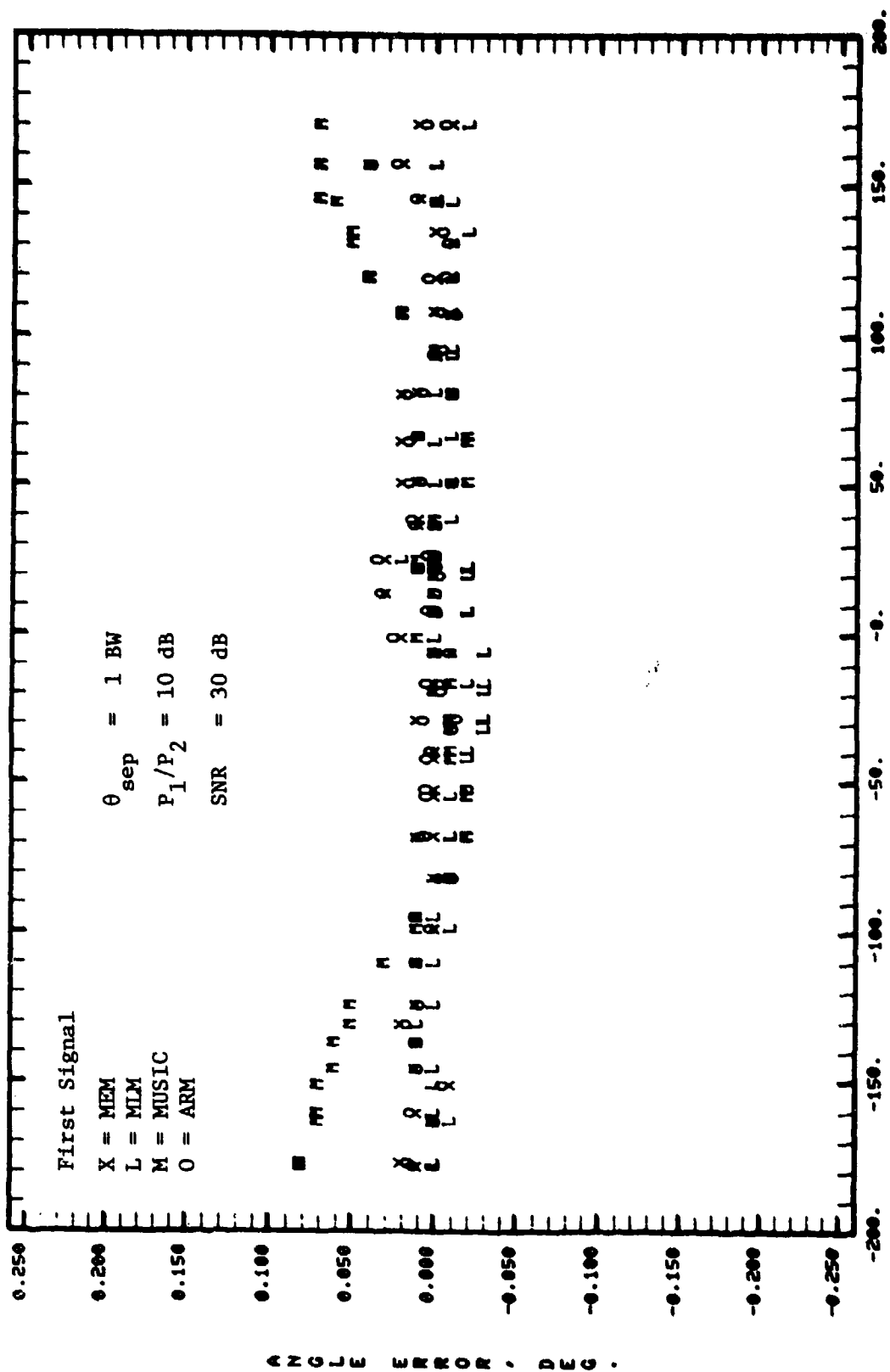
Fig. 5-38. Angle errors: first signal in Case 2.

First Signal

X = MEM
L = MLM
M = MUSIC
O = ARM

θ_{sep} = 1 BW

P_1/P_2 = 10 dB

SNR = 30 dB
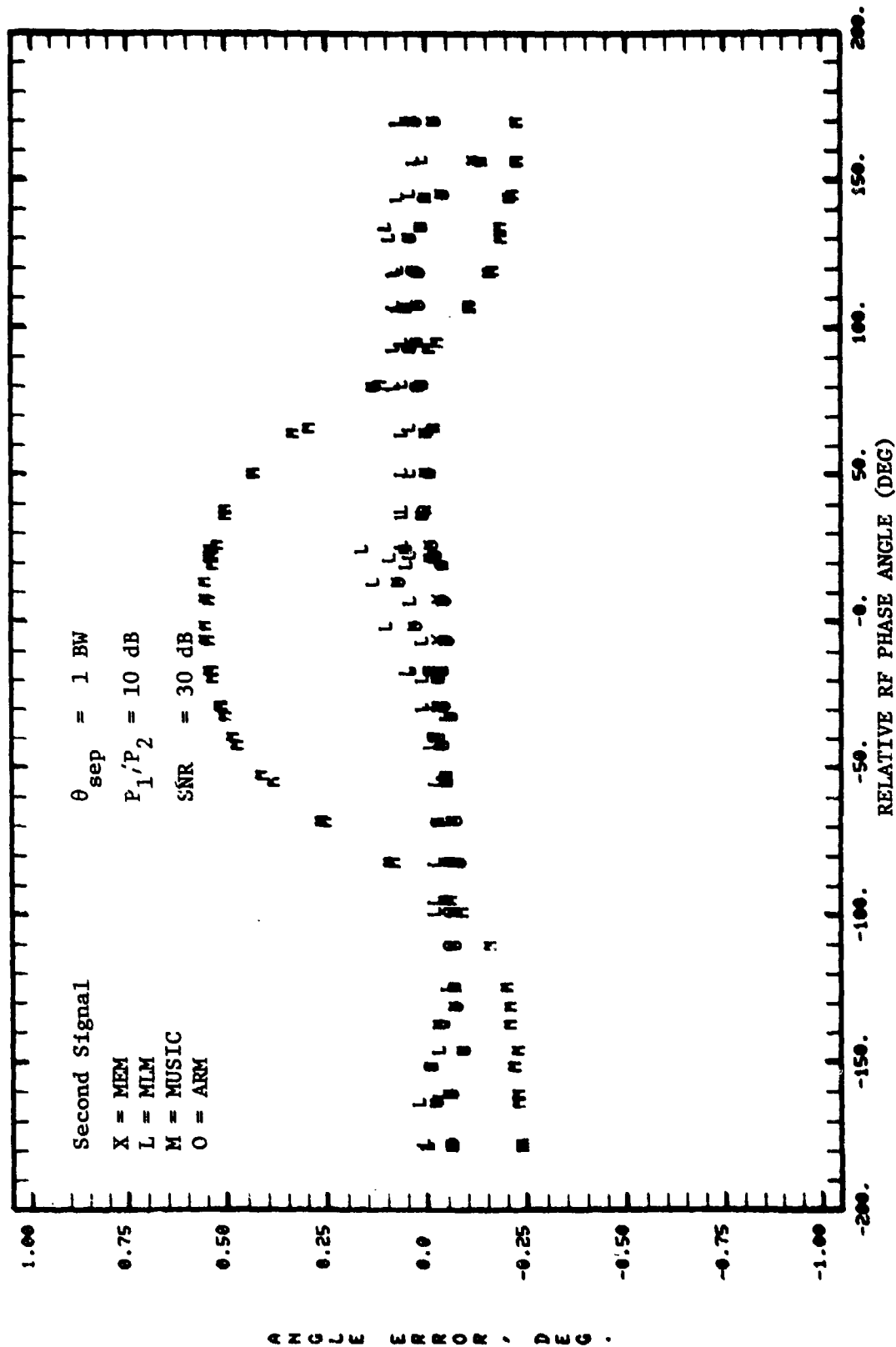
Fig. 5-39. Angle errors: second signal in Case 2.

5-60
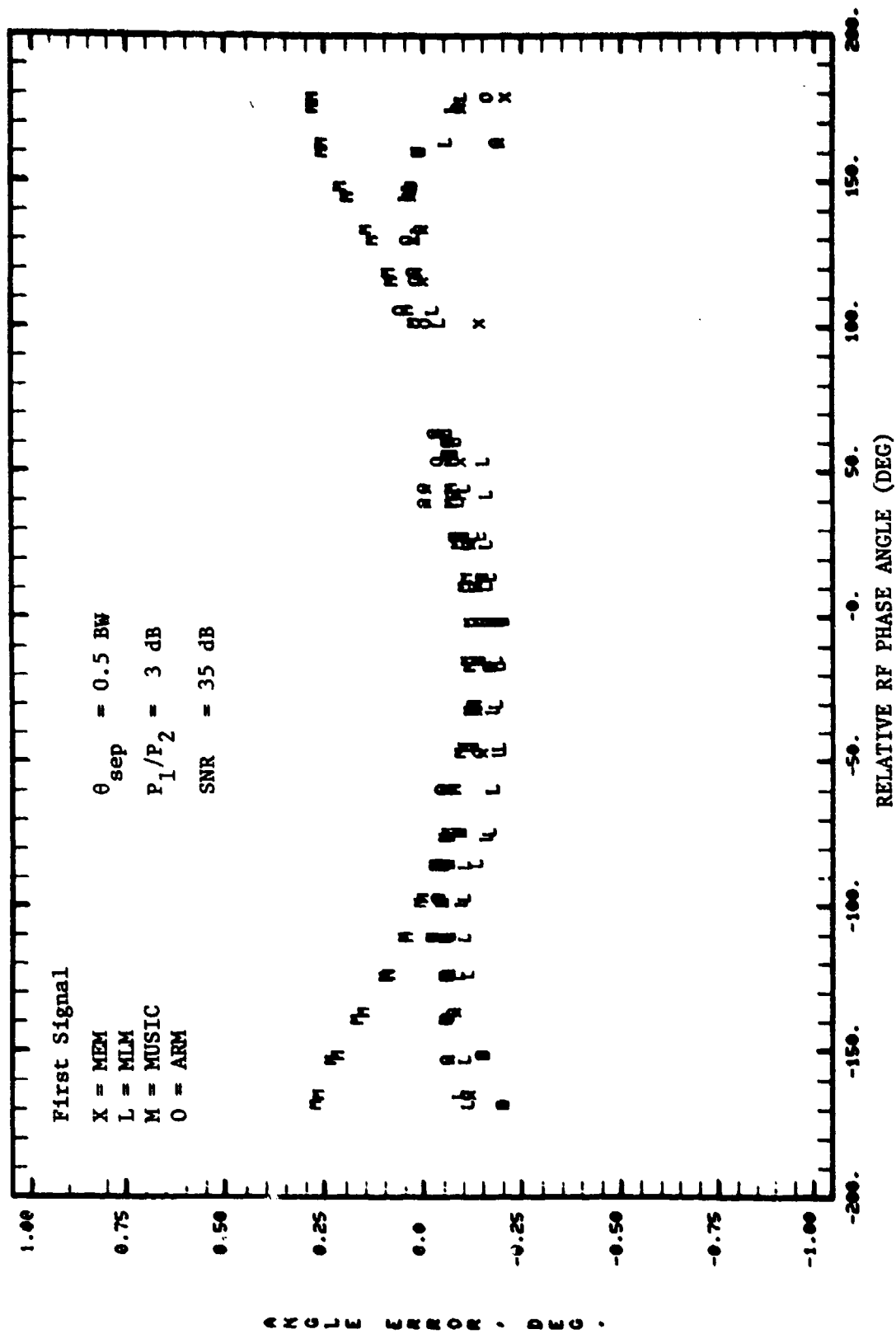
Fig. 5-40. Angle errors: first signal in Case 3.
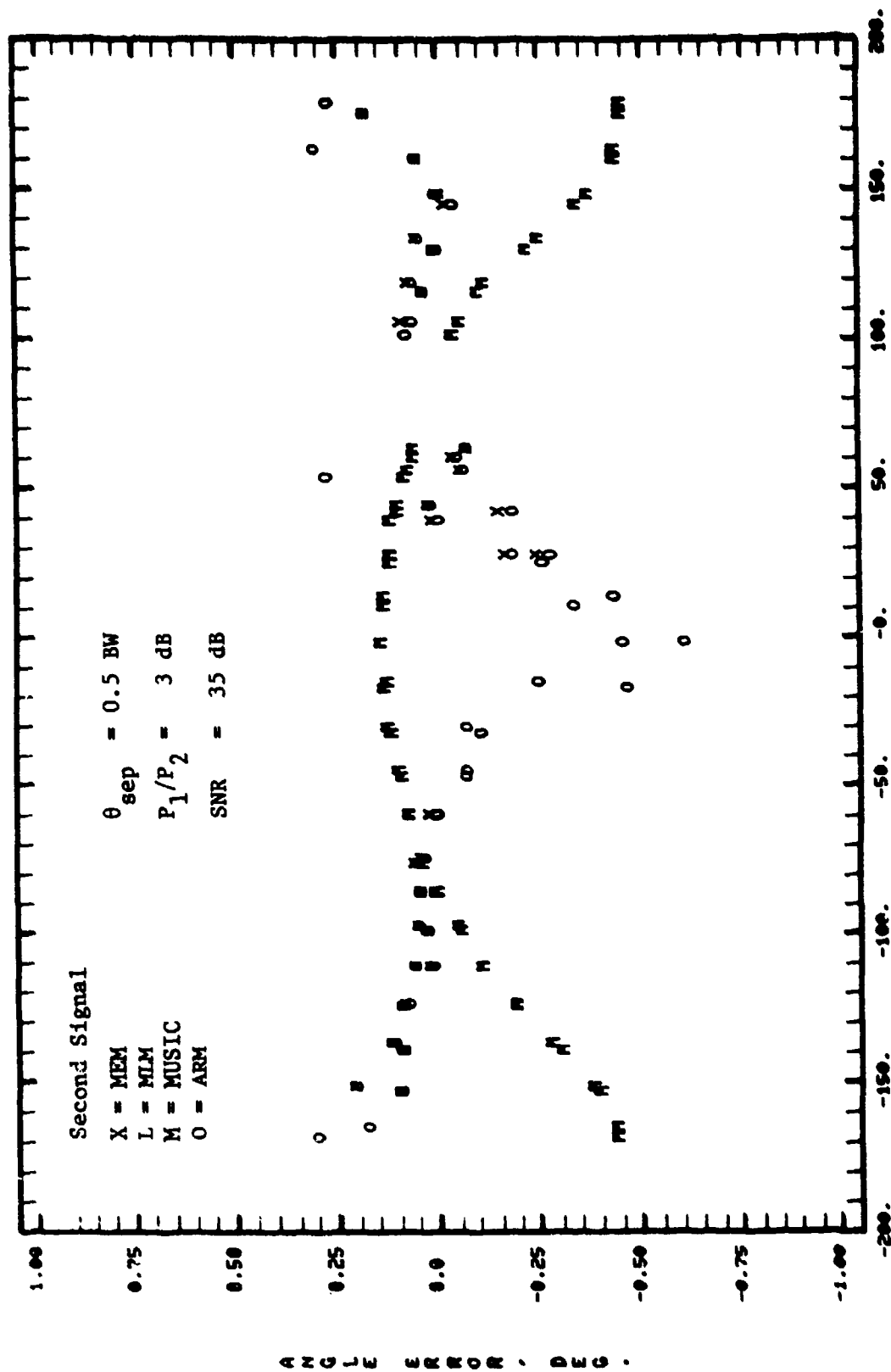
Fig. 5-41.   Angle errors: second signal in Case 3.

5-62
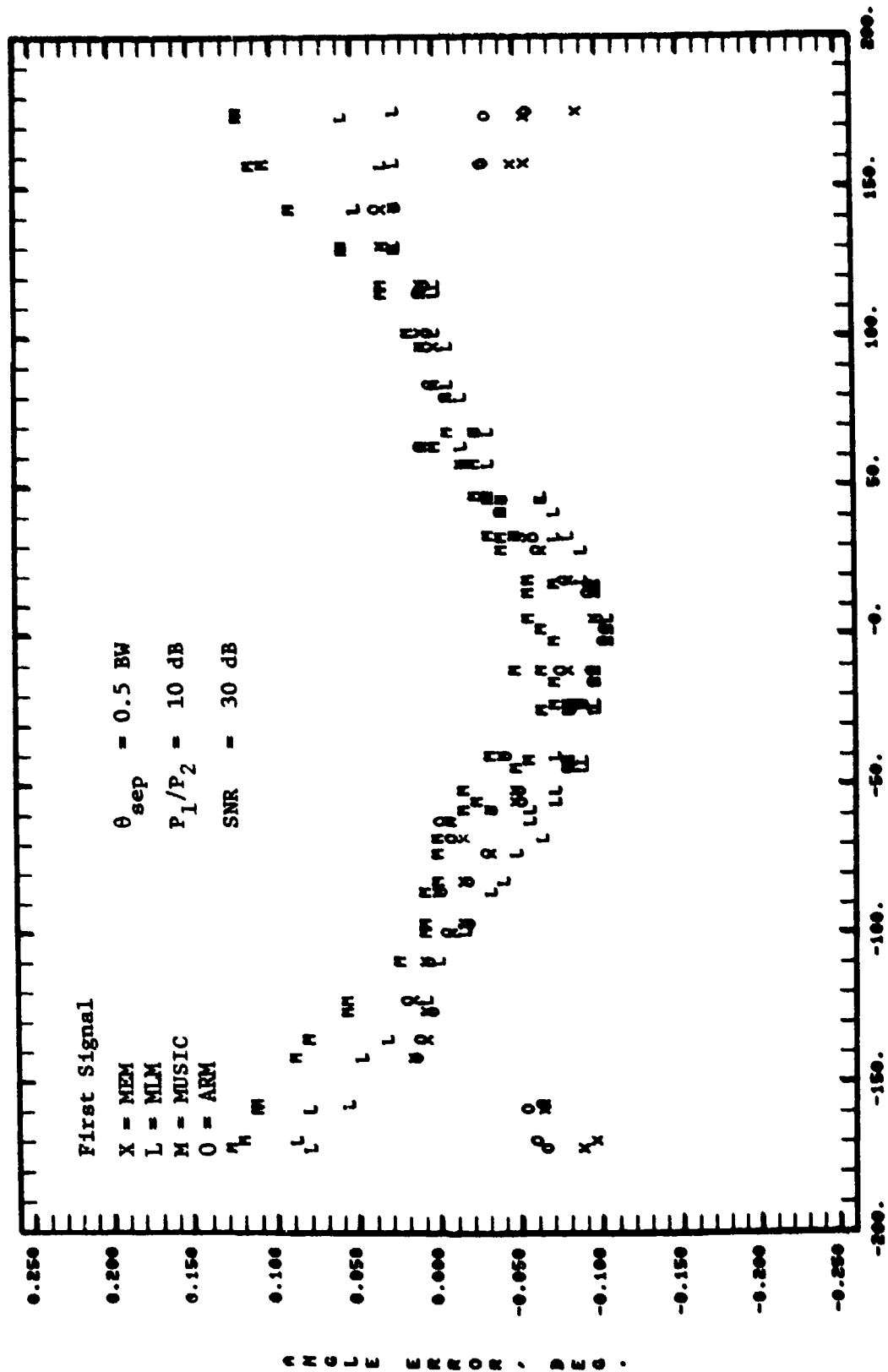
Fig. 5-42. Angle errors: first signal in Case 4.
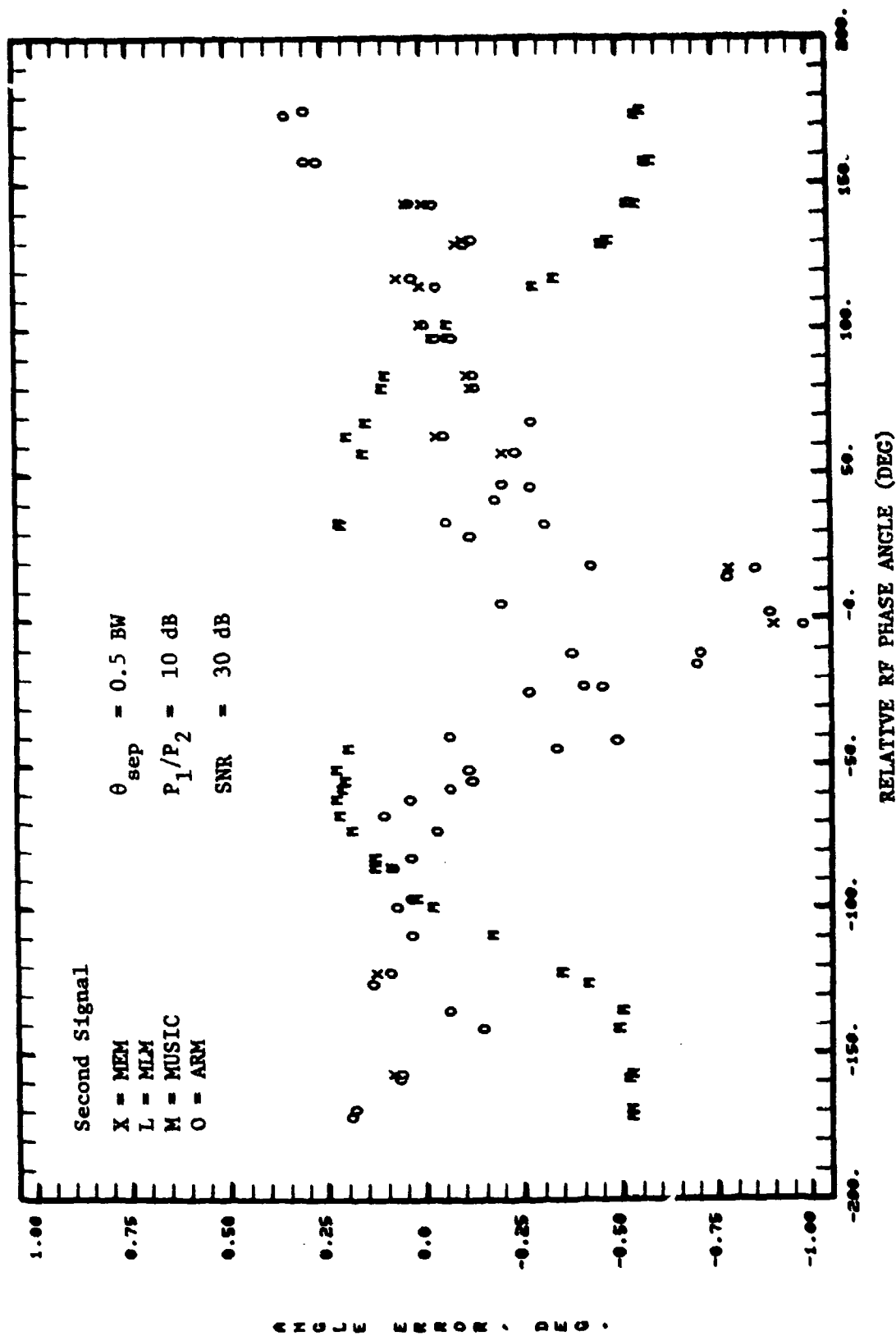
Fig. 5-43. Angle errors: second signal in Case 4.

5-64

nal. For this case, only ARM indicates the existence of two signal components in the incoming signal for all relative phases between two signals. MUSIC misses the detection of the second signal around $0°$ relative phase, while the MEM yields two signal detections only for the relative phases around $\pm90°$. Again, MLM fails to reveal the second signal. For the first signal, the peak ARM angle errors are about 0.09 BW for the relative phases near $\pm180°$ and $0°$. For the second signal, the angle errors are much larger, especially near the $0°$ relative phase where errors in the order of $0.7 \sim 1.0$ BW are observed.

Table 5-4 summarizes the results for the above four cases. For the larger angular separation between two signals such as 1 BW in the first two cases, it appears that (1) all four processing techniques can resolve two signals, and (2) the performance of the MUSIC technique was poorer than the other three techniques in terms of angle error and much more sensitive to the relative phase between the two signals. For the smaller angular separation, such as 0.5 BW in the last two cases, only the ARM technique can detect two signal sources for all relative phases; however, the angle errors are considerably greater than those in the first two cases of larger angular separation.

It should be noted, however, that there was only a single true snapshot of array data available at each relative phase angle. The Monte Carlo and deterministic simulation results suggest that the eigenanalysis and MEM/ARM techniques would have shown more sizable performance advantages had some degree of time averaging been utilized.

2. Second Group of Field Data: Unknown Multipath Environment

The bulk of the discussion and analysis above has centered on detection and parameter estimation for a small number of plane waves. In this section, we describe the application of the high-resolution techniques to some experimental terrain scattering field data in which our objective is to characterize the characteristics of reflection environment. The approach to performance assessment here will differ considerably from that in the earlier portion of this chapter in that the "true" number and amplitudes of planes waves is not known. Further complicating the data in this case is the possibility of diffuse multipath, i.e., signals arriving from an extended angular region.

5-65

TABLE 5-4

SUMMARY OF RESULTS FOR TWO SIGNAL CASES

	Technique	Detection of Both Two Signals[*]	$\Delta\theta$[**] (BW)		Strong Dependence of Angle Error on ϕ[***]
			1st Signal	2nd Signal	
Case 1	MEM	1	0.07	0.1	No
	MLM	1	0.04	0.1	No
	MUSIC	1	0.1	0.4	Yes
	ARM	1	0.07	0.1	No
Case 2	MEM	1	0.03	0.1	No
	MLM	1	0.03	0.3	No
	MUSIC	1	0.08	0.5	Yes
	ARM	1	0.04	0.1	No
Case 3	MEM	1/2	0.2	–	Yes
	MLM	0	–	–	Yes
	MUSIC	1	0.3	0.45	Yes
	ARM	1	0.2	0.5	Yes
Case 4	MEM	1/2	0.1	–	Yes
	MLM	0	–	–	Yes
	MUSIC	1/2	0.13	–	Yes
	ARM	1	0.1	1.0	Yes

[*]Detection of both two signals, 1: for all ϕ, 1/2: for part of ϕ, 0: no detection

[**]$\Delta\theta$: Peak angle error in beamwidths (BW)

[***]ϕ: Relative phase between two signals .

Case 1: θ_{sep} = 1 BW, $P_1/P_2 \approx$ 3 dB, SNR $\approx$ 33 dB

Case 2: θ_{sep} = 1 BW, $P_1/P_2 \approx$ 10 dB, SNR $\approx$ 30 dB

Case 3: θ_{sep} = 0.5 BW, $P_1/P_2 \approx$ 3 dB, SNR $\approx$ 35 dB

Case 4: θ_{sep} = 0.5 BW, $P_1/P_2 \approx$ 10 dB, SNR $\approx$ 30 dB

Spectrum shape is viewed as a possibly important parameter if one is to distinguish between a discrete specular reflection environment as opposed to a mixture of specular and diffuse multipath.

Figure 5-44 shows the equipment setup for the ground-reflection field measurements to obtain this group of field data. Signals (1090-MHz L-band and 5.2-GHz C-band) were radiated from a helicopter, flying vertical descent/ascent flight profiles at distances of 0.5 to 1.5 nmi from the receiving antenna arrays. The incoming signal at the receiving antenna will normally consist of a signal arriving at a positive elevation angle (corresponding to the direct signal coming from the helicopter) and other signals generally at negative elevation angles (corresponding to various ground reflections from terrain features.) Here our objective is to characterize the elevation multipath environment (i.e., number of multipath components, their arrival angles and signal levels) through the angular power spectral estimates of the incoming signal by various signal processing techniques. The results shown below were obtained with a 9-element 25.9λ-aperture L-band array and a 19-element 56.7λ aperture C-band array.

Figure 5-45 shows the L-band angular power spectral estimate for a flight test in which the target helicopter was at an elevation angle near $2°$. Figure 5-46 gives the corresponding C-band result. The terrain in front of the receiving antenna array consisted of a fairly flat grass field adjacent to the main runway at Hanscom Airport, Massachusetts. For the L-band data, it can be seen that all four techniques suggest the presence of two signals, one direct signal and one ground-reflected signal. All four techniques yield the same angle and relative power[*] estimates of two arrival signals. In terms of the appearance of the angular power spectrum, it appears that the MEM shows sharp-

---

[*]The estimation of the power relative to the direct signal was obtained by least-squares fit of the measured data with the signal model consisting of various plane waves arriving at the estimated angles. The direct signal power is taken to be 0 dB. The relative power levels are indicated by the letters 'X', 'L', 'M', and 'O' for MEM, MLM, MUSIC, and ARM, respectively.
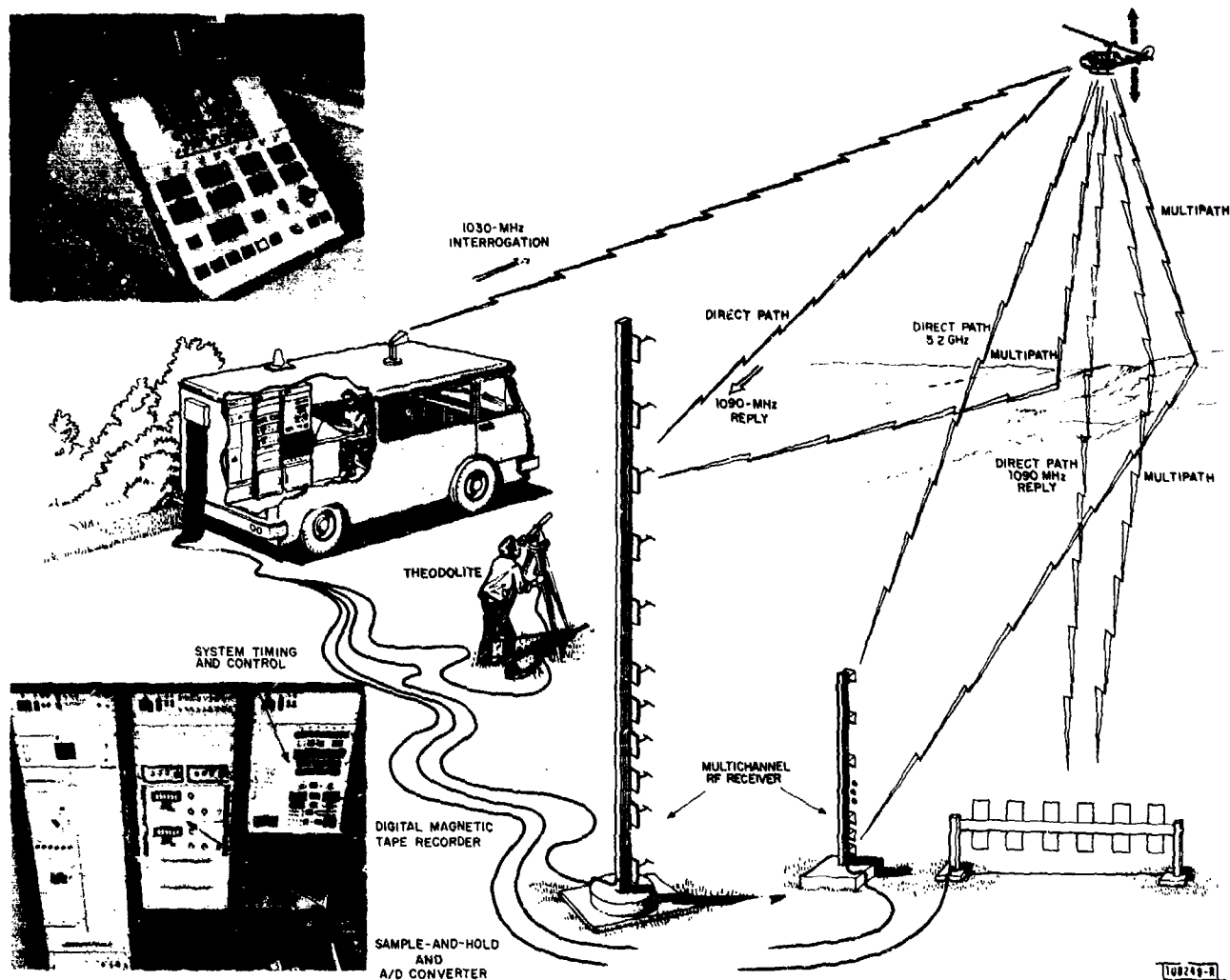
Fig. 5-44. Aperture sampling experimental con-
figuration for terrain multipath measurement.

L-Band Data

X : MEM
L : MLM
M : MUSIC
O : ARM

Theodolite Angle

ANGLE IN DEG

POWER IN DB
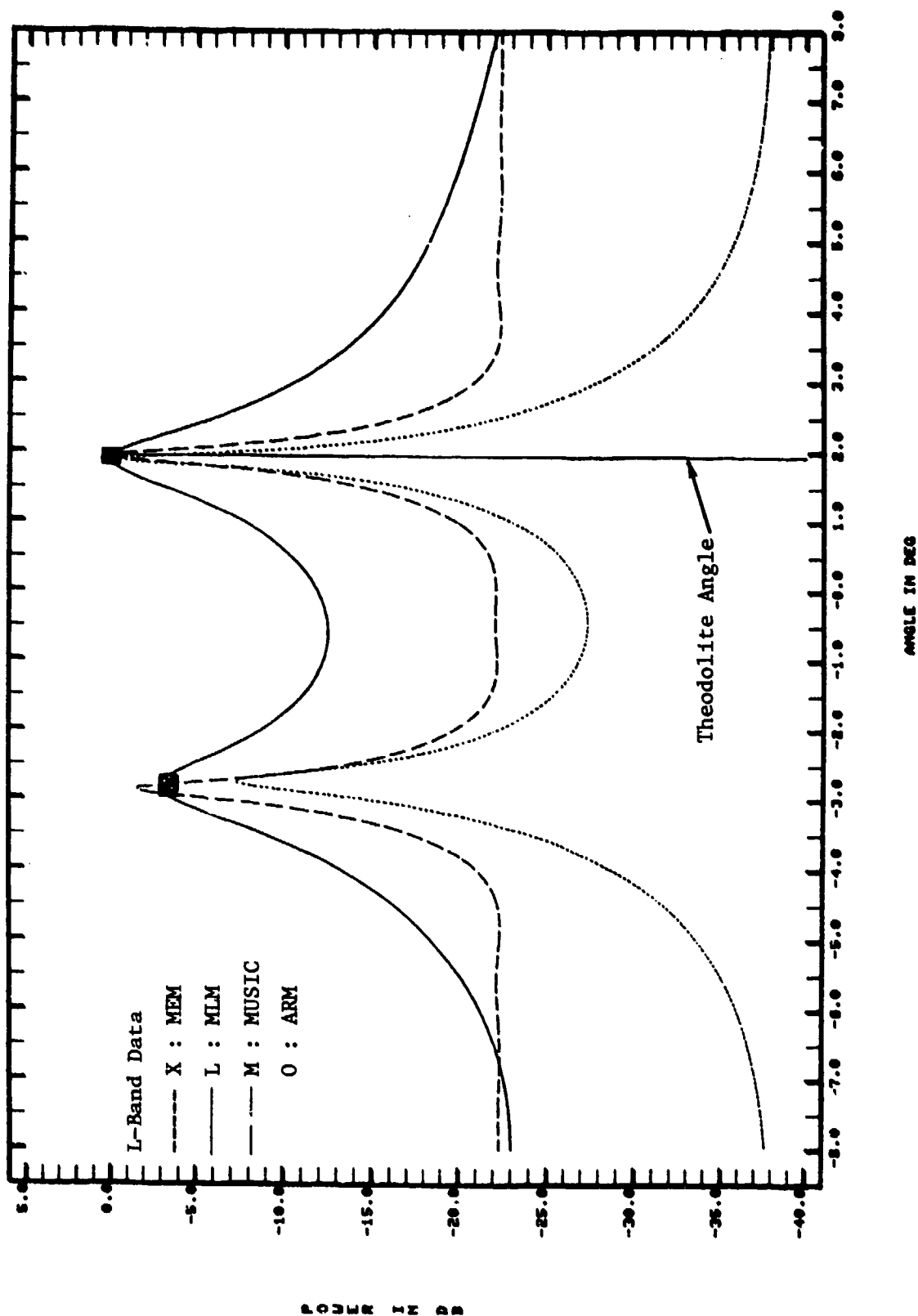
Fig. 5-45.   Angular power spectral estimates: Hanscom Airport L-band measurement.
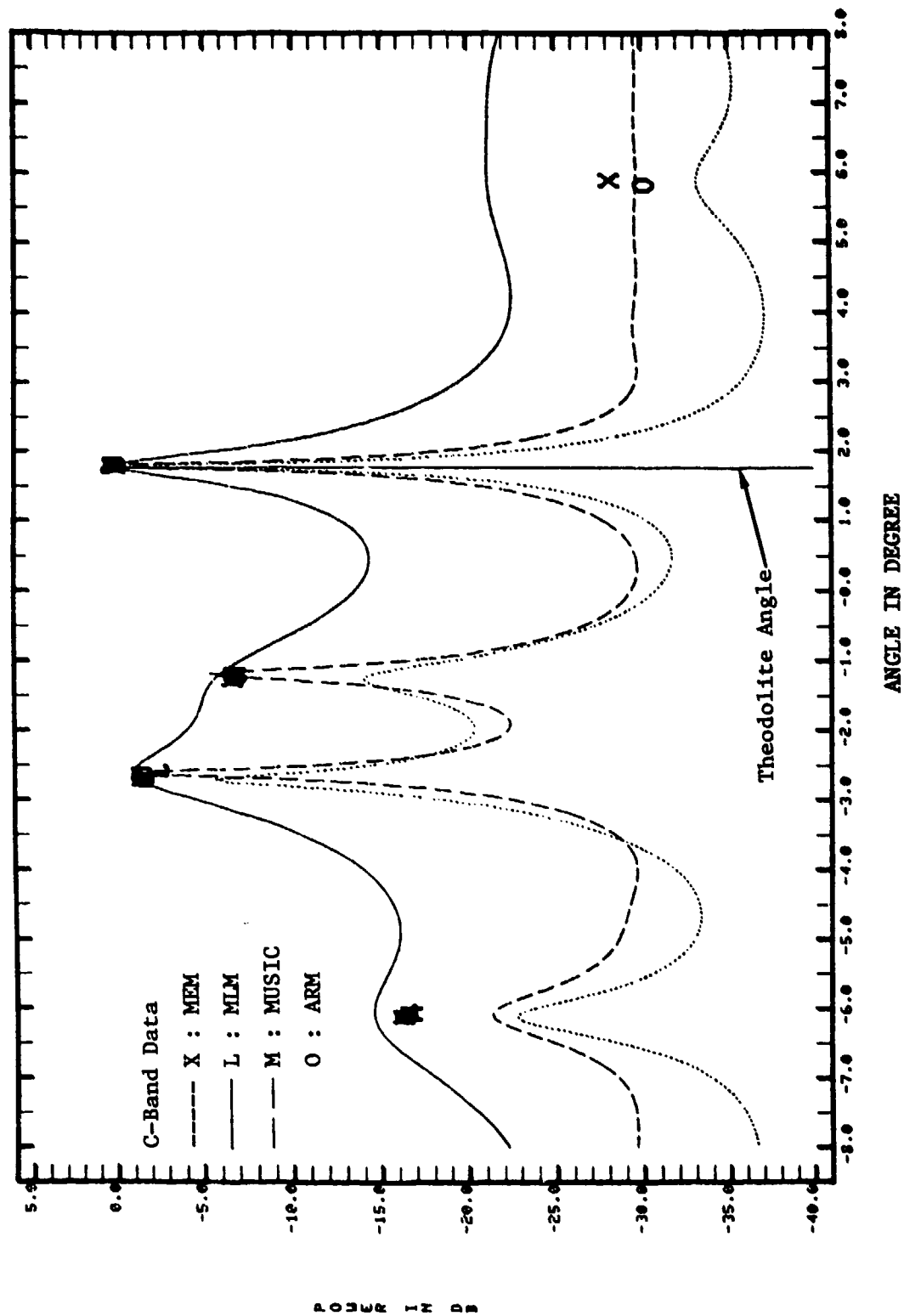
Fig. 5-46. Angular power spectral estimates: Hanscom Airport C-band measurement.

er spectral peaks and much lower background spectral level. For the C-band data, ignoring the spectral peak around -30 dB, again all four techniques yield the same number of signal arrivals (4), the direct signal at 1.8°, and the other three signals at negative elevation angles, which are believed to be the ground-reflected signals. Also, the indicated signal levels and arrival angles of various signal components are very much the same among the four techniques, except that the identification of various reflected signals with the MLM technique is more difficult than as that with other techniques.

Figures 5-47 (L-band) and 5-48 (C-band) show the spectral estimates for a flight test in which the target helicopter elevation angle was around 4.2°. This field test was taken at the golf course at Fort Devens, Massachusetts, where the terrain was rolling. For the L-band data, all four techniques suggest the arrival of two ground-reflected signals and yield very similar angle and power estimates of these signal components. A compared to the MEM and MUSIC spectral estimates, the MLM angular power spectrum has much broader peaks and much higher background spectral level. For the C-band data, the four techniques do not quite agree with each other in terms of the number of signal arrivals. However, for those three signal components which appear in all four estimates, all four techniques give the similar indication of arrival angles and signal levels. ARM and MUSIC indicate one more signal arrival around 5.5°[*] at -6 dB ~ -8 dB, while ARM and MEM suggest another signal arrival at -0.6° at -15 dB.

In the four cases shown above (Figs. 5-45 to 5-48), we notice that the multipath separation angles (i.e., the separation angle between the direct signal and the closest ground-reflected signal) are all more than two beamwidths. For this larger multipath separation, all four processing techniques accurately estimated the target helicopter elevation angle[**] and gave

_____

[*]If we interpret this signal arrival to be a ground reflection, the true arrival angle should be around -12° (since the grating lobe of our C-band array appears at about every 18°).

[**]The larger angle deviation of the estimated angle from the theodolite tracking angle observed in Fig. 5-47 is believed to be caused by erroneous theodolite tracking.
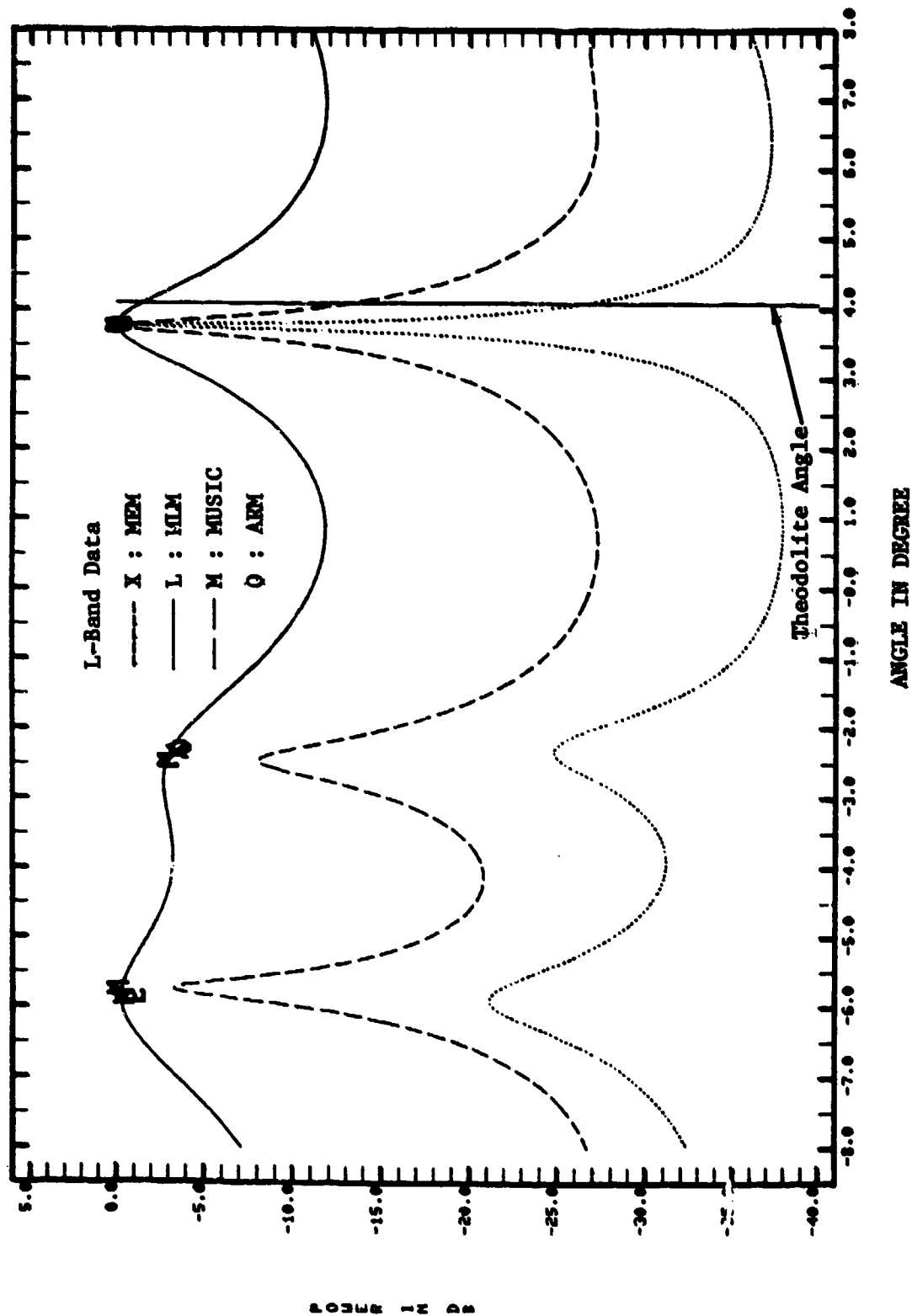
Fig. 5-47. Angular power spectral estimates: Fort Devens L-band measurement.

ANGLE IN DEGREE
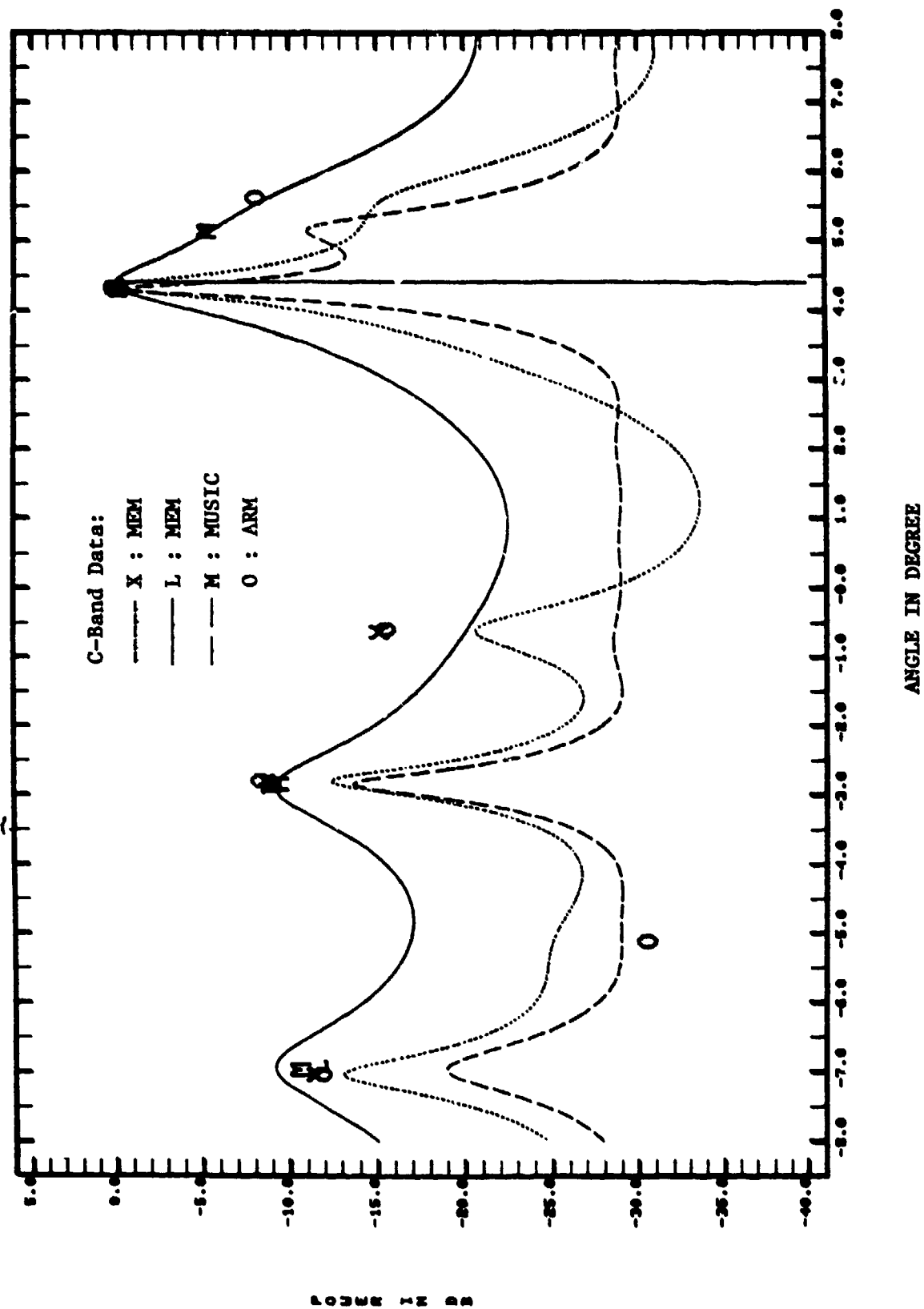
Fig. 5-48. Angular power spectral estimates: Fort Devens C-band measurement.

5-73

essentially identical estimates of the arrival angles and signal levels of major ground reflected signal components. The multipath separation angles for the following two cases are less than one beamwidth.

Figures 5-49 (L-band) and 5-50 (C-band) show the spectral estimates for a flight test at Otis AFB, Massachusetts, in which the target helicopter was at around 0.5° elevation. The terrain at this test site consisted of various small-scale bumps on a fairly flat piece of barren ground. For the L-band data, the MEM, ARM, and MUSIC indicate one high-level ground-reflected signal in addition to the direct signal arrival at 0.6°, while the MLM only shows a broad peak at 0°. For the C-band data, although MEM, ARM, and MUSIC again suggest one strong ground-reflected signal, the estimates of the arrival angle and the multipath level are somewhat different (-0.5° at -2 dB indicated by the MEM and ARM versus -0.4° at +2 dB given by the MUSIC). Also, as previously observed, the indications of various low-level signal arrivals are different among various processing techniques. Again, as observed in the L-band data, MLM only gives one spectral peak at 0°.

For these two cases, the separation angles between the direct and the ground-reflected signals are about 1.2° (or 0.6 BW) and 0.8° (or 0.8 BW) for the L- and the C-band cases, respectively. This gives us the opportunity to compare results of terrain reflection measurements here with those of two signal measurements at the antenna test range discussed previously. To make the comparison, we need to know the RF phase of the ground-reflected signal relative to the direct signal at the phase center of our L-band (or C-band) antenna array. This RF phase[*] of the ground reflection is estimated to be about -50° for the L-band case and about -160° for the C-band case. For the L-band case (0.6 BW separation and -50° relative phase), the two signal measurement results for the case of 0.5 BW separation and 3 dB power ratio

---

[*] The RF phase of the ground-reflected signal relative to the direct signal at the phase center of the receiving array was estimated assuming the helicopter was above the flat ground at the theodolite tracking angle. The RF phase estimate includes the phase delay due to the travel path difference between the direct signal and the ground reflection as well as the phase change due to the reflection on the ground.
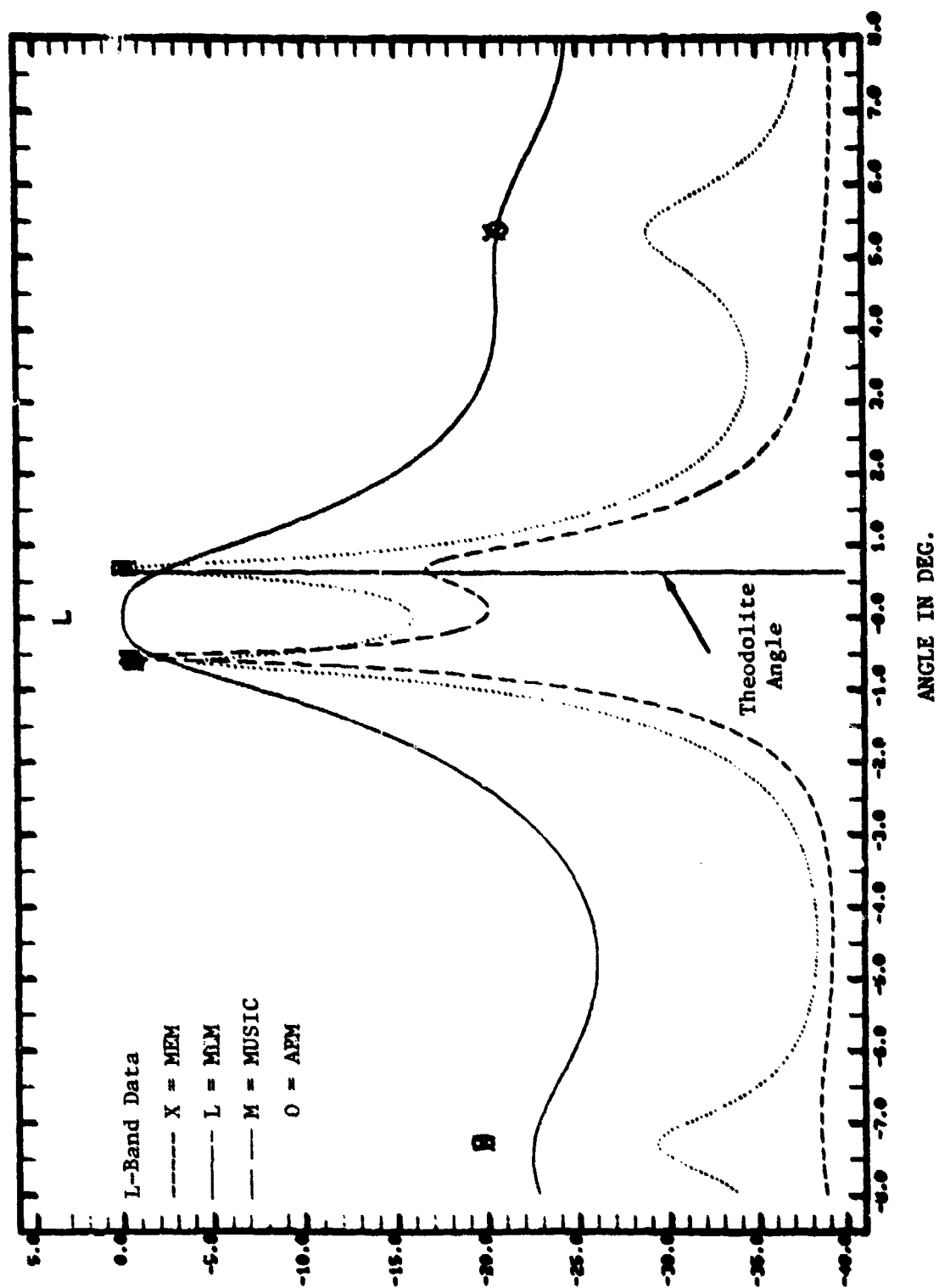
Fig. 5-49. Angular power spectral estimates: Otis AFB L-band measurement.

L-Band Data

X = MEM
L = MLM
M = MUSIC
O = APM

Theodolite
Angle

ANGLE IN DEG.
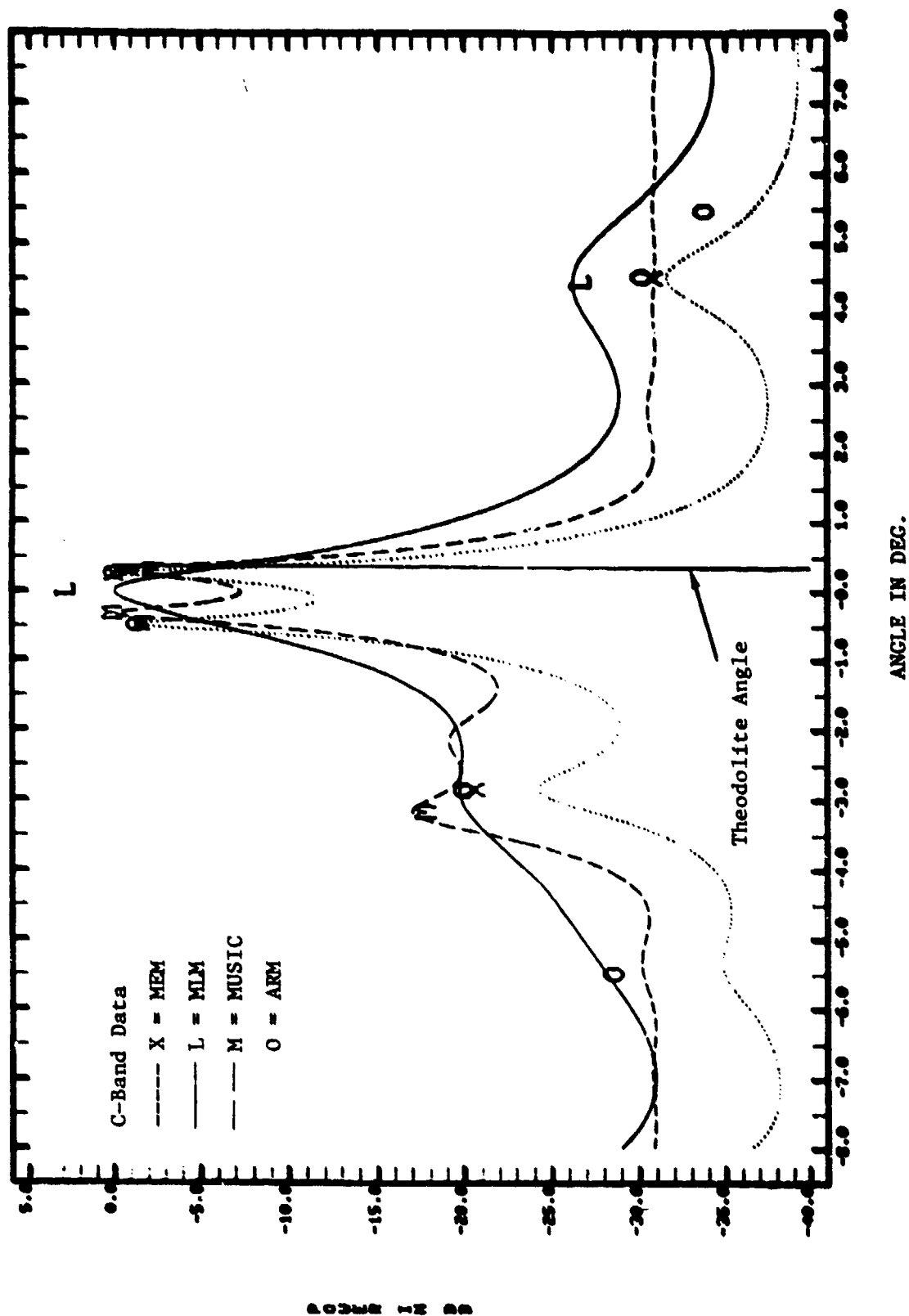
Fig. 5-50. Angular power spectral estimates: Otis AFB C-band measurement.

(Figs. 5-40 and 5-41) suggest that (1) the MLM technique cannot resolve two signals and (2) the MEM, ARM, and MUSIC techniques should give the fairly accurate angle estimates for both signal arrivals. This agrees very well with that being observed here in the L-band spectral estimates (Fig. 5-49). For the C-band case (0.8 BW separation and -160° relative phase), the two signal measurement results for the case of 1 BW separation and 3 dB power ratio (Figs. 5-36 and 5-37) suggest that all four processing techniques can resolve the two signals and (2) MEM, ARM and MLM techniques should give better angle estimates than the MUSIC technique. This also agrees fairly well with that being observed in the C-band spectral estimates (Fig. 5-50), except that in Fig. 5-50 the MLM technique fails to resolve the two signals. The disagreement in the resolution capability of the MLM technique between the two signal measurement results and the C-band results shown here probably is due to the fact that the separation angle is 0.8 BW for the C-band case here instead of the larger separation angle of 1 BW in the two signal measurement results in Figs. 5-36 and 5-37.

For terrain multipath field data, it appeared that all four signal processing techniques would give similar results for a simple terrain condition such as the near-flat terrain for the L-band data taken at Hanscom Airport. However, when the multipath environment became more complicated, such as in the rolling terrain or for the C-band data, the results from four techniques often only agree on the signal arrivals with larger power. In general, the identification of various signal arrivals with the MLM estimate was not as clear as the other estimates, especially when the multipath separation angle is less than one beamwidth.

VI. HARDWARE IMPLEMENTATION ISSUES

In this chapter, we consider some of the hardware implementation issues associated with use of the advanced array processing techniques discussed in the preceding chapters. The focus here will be on those issues which emerged as important in the sampled-aperture system shown in Fig. 6-1. This system has been sucessfully used for several years in terrain multipath and plane wave resolution experimental studies.

The block diagram shown in Fig. 6-2 shows the current MLS studies sampled-aperture measurement system. The various signal processing techniques are implemented in software on a general-purpose computer (both on a Data General Eclipse minicomputer and on the Lincoln timeshare system). The sampled-aperture field data are obtained from linear antenna arrays together with a multichannel receiver and a digital data acquisition system. Since a general-purpose computer with double-precision floating point was used, no computation problems arose in the signal processing part of the system after the software had been implemented and debugged. Thus, our discussion will focus on the aperture-sampling part of the measurement system, i.e., the first block in Fig. 6-2.

First, we will describe those problems which are commonly encountered in the hardware implementation of a multichannel system for aperture sampling and the methods which employed in the existing system to overcome those problems. Then, we will discuss the system calibration approach used in this equipment and the principal problems encountered during several years of field measurements. The final section comments on some of the issues involved in developing an operational real-time aperture-sampling system.

A. Generic Features

The principal end product of the measurement system is good-quality sampled-aperture data, i.e., the correct amplitude and phase values of the incoming wavefront at the various antenna elements. For an aperture-sampling system using with several multichannel subsystems as shown in Fig. 6-2, the quality of the measured data principally depends on the channel tracking or

6-1
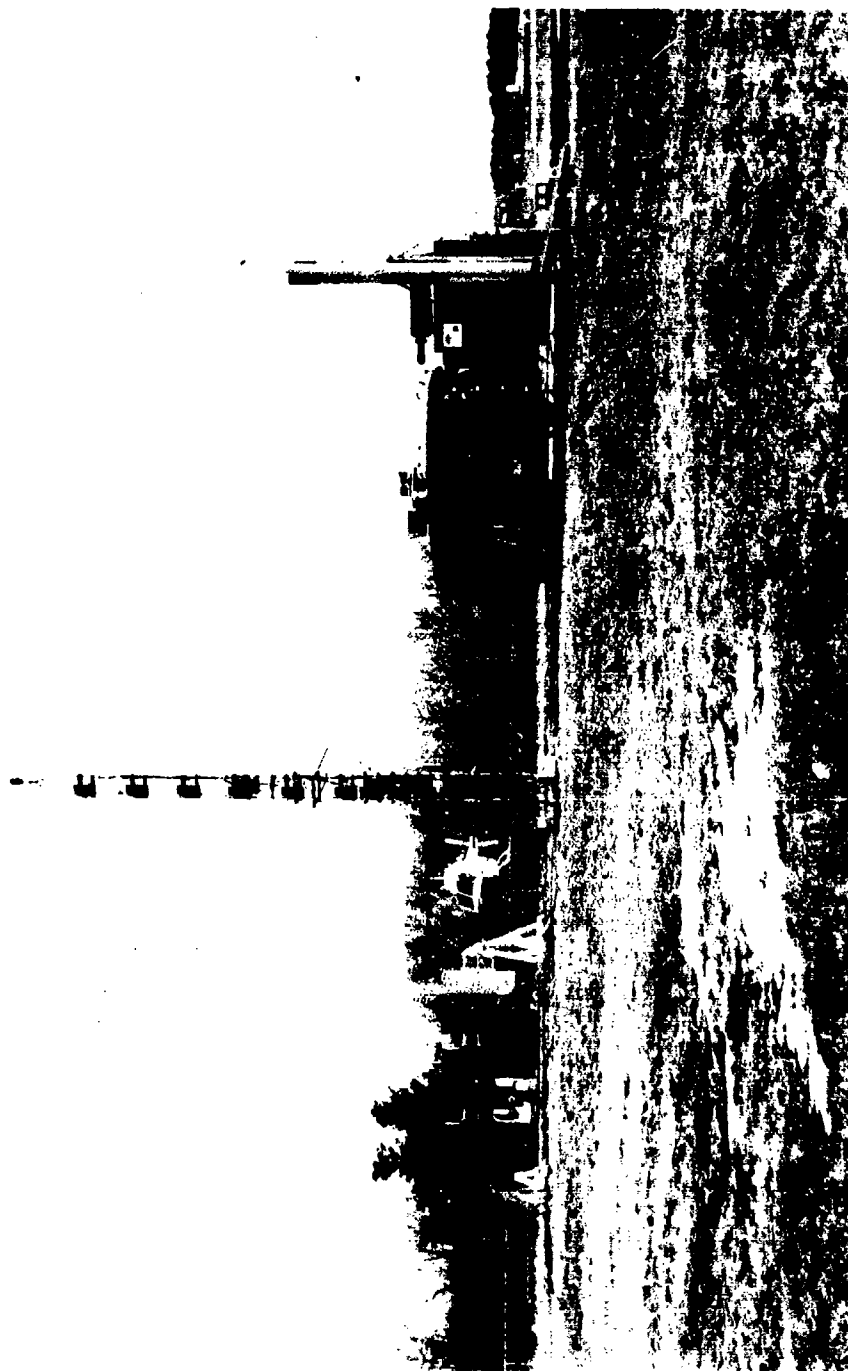
Fig. 6-1. Experimental aperture sampling system in terrain multipath measurement mode.
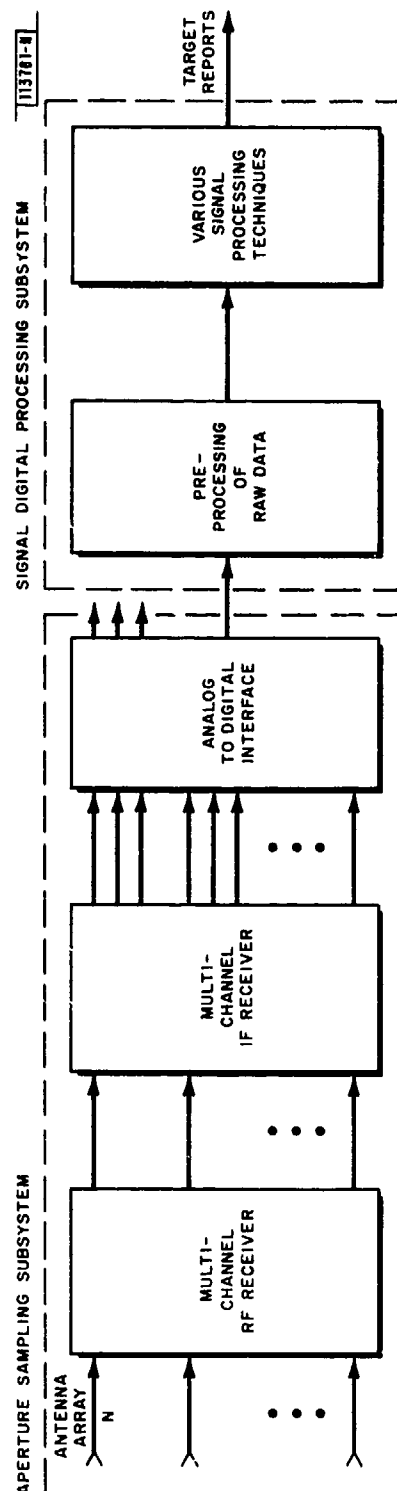
Fig. 6-2. Simplified block diagram of Advanced Array Processing System.

equalization. The main contributions to the channel mistracking are the channel amplitude and phase biases due to either the physical component dissimilarity among channels and/or differential channel sensitivity to environmental factors such as temperature and humidity.

For an aperture-sampling system using an N-element antenna array feeding an N-channel RF/IF receiver, each channel typically will consist of an antenna element, an RF front end which is formed by a preamp/mixer and some RF filters, and the IF section which consists of filters, amplitude and phase detectors. The various components in each channel are normally connected together through various RF and IF cables. Of course, one would like to have a system built with all channels being exactly identical. However, more realistically, there will be variations among all channels in their amplitude and phase characteristics due to component differences. Also, the transfer function of various components in a given channel can be sensitive to the changes in the temperature and/or humidity. The temperature variation in the component behavior, particularly that of the components in the RF front end and the RF cables, was found to be a major factor in channel tracking for the system shown in Fig. 6-1.

When the number of the antenna elements is greater than the number of the available receiver channels and the required data rate is slow, RF switches can be used to serve the time sharing or channel-multiplexing purpose. For example, five antenna elements can share the same receiver channel by using a single-pole 5-throw (SP5T) RF switch. A RF switch is a multichannel device by itself, and thus additional channel tracking issues will arise although use of a common IF system reduces certain other problems.

The conversion to digitized aperture sample data is accomplished by an analog-to-digital (A/D) converter at the output of the IF amplitude and phase detectors. The error in the A/D conversion is another contributing factor, in addition to the channel tracking error, to the total measurement error in the sensor signal.

6-4

B.   Experience with The Current Measurement System

Figure 6-3 shows a more detailed block diagram of the current aperture-sampling system with two 5-channel RF receivers. One receiver operates at 1090 MHz (L-band) with 17 antenna elements, and the other operates at 5.2 GHz (C-band) with 29 horn antennas. Both the L- and C-band RF receivers share the same 5-channel IF receiver. A single time multiplexed 8-bit A/D converter serves all five IF channels.

As previously mentioned, the quality of measured data produced by an aperture-sampling system is a strong function of the channel tracking. To minimize channel tracking problems, our approach was to build equalized antenna/receiver channels to start with and, to make various system calibrations before and after measurements to remove any residual channel amplitude and phase biases.

1.   Design Features to Yield Good Stability

Here, we will discuss those features which we implemented in the system hardware to achieve good channel equalization and differential stability. For this discussion, we will separately consider three parts - the L-band subsystem, the C-band subsystem, and the common IF subsystem. Referring to Fig. 6-3, the L-band subsystem covers the system path from the L-band antenna element down to the L-band IF bandpass (BP) filter, the C-band subsystem is from the C-band antenna element down to the C-band IF BP filter, and the IF subsystem is the common phase detectors and log amplifiers shared by both the L-band and C-band subsystems.

a.   L-band Subsystem

This was built around the existing 5-channel RF front ends from the Precision and Approach Landing Monitor (PALM) system [42], with a (new) elevation array consisting of eleven new dipole antennas and a (new) azimuth array consisting of six original PALM antenna elements. The eleven dipole antennas are physically identical and have a close electrical match (E-plane gain = 8.3 ± 0.7 dB and VSWR = 1.5 ± 0.09 : 1). The original PALM antenna elements were also designed to be identical.

6-5
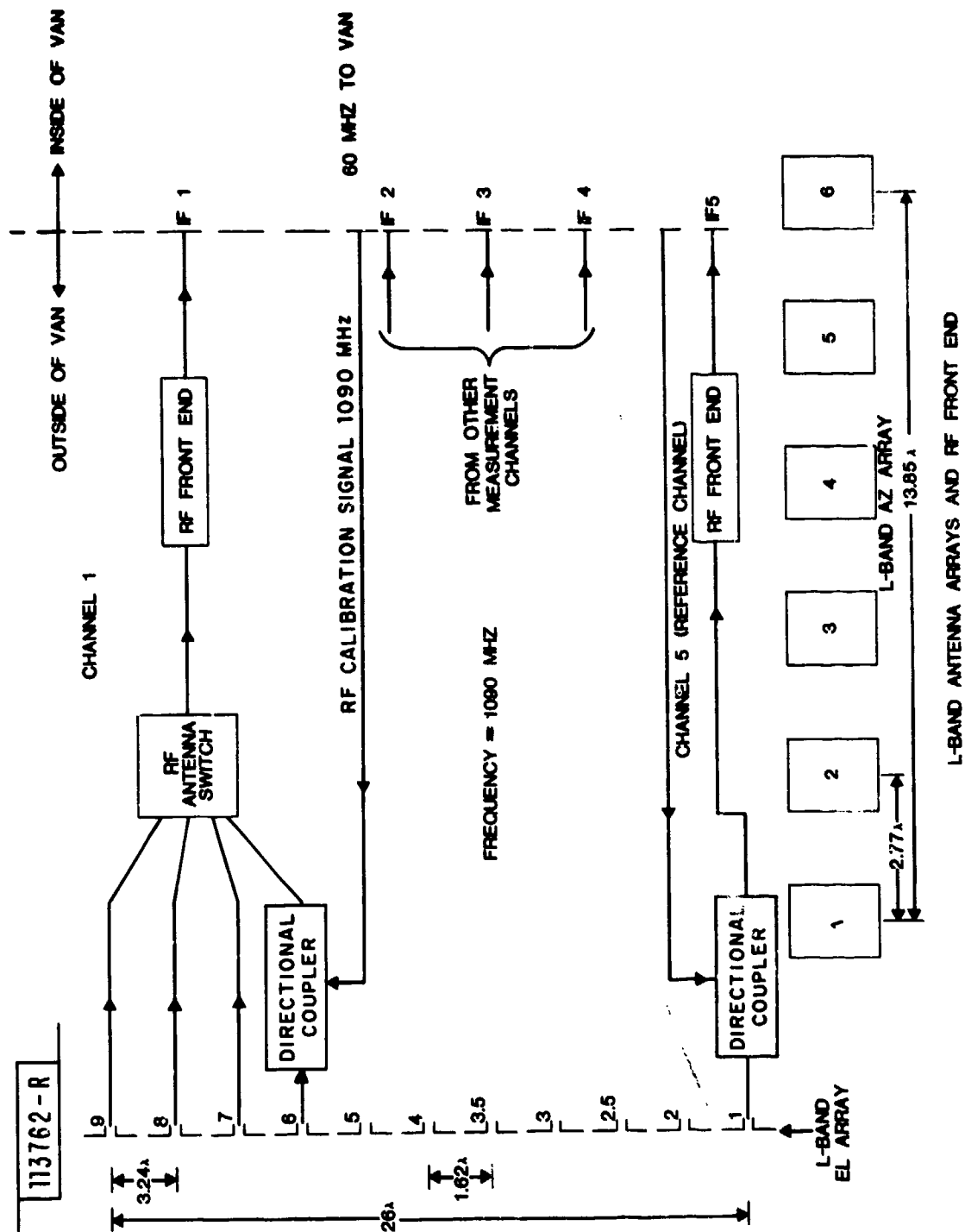
Fig. 6-3a. Block diagram of current aperture sampling system: L-band antenna arrays and RF front end.

C-BAND ANTENNA ASSEMBLY

WAVE GUIDE ASSEMBLY

60 MHZ TO VAN

60 MHZ TO VAN

60 MHZ TO VAN

60 MHZ TO VAN

60 MHZ REFERENCE CHANNEL TO VAN

F1

F2

F3

F4

F5

5.222 GHZ OSCILLATOR

5.162 GHZ TEST OSCILLATOR

PREAMP

TEST OUTPUT

8 WAY DIVIDER

IF

7 DBM LO

SP10T

RF CALIBRATION SIGNAL

3.15 λ

1.57 λ

1 2 3 4 5 6 7 8 9 10 11 12 13 14 15 16 17 18 19 20 21 22 23 24 25 26 27 28 29 30
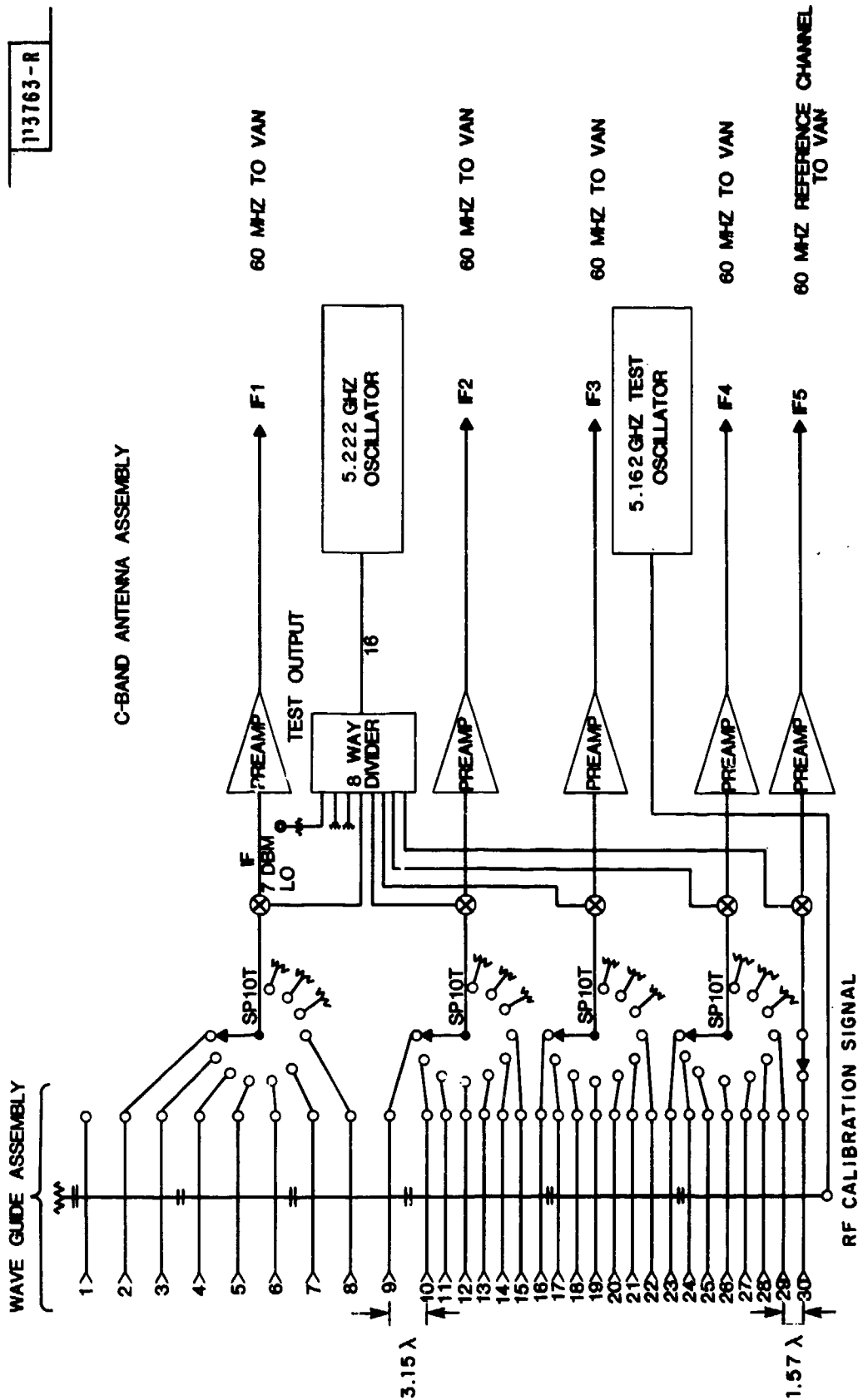
Fig. 6-3b. Block diagram of current aperture sampling system: C-band antenna array and RF front end.
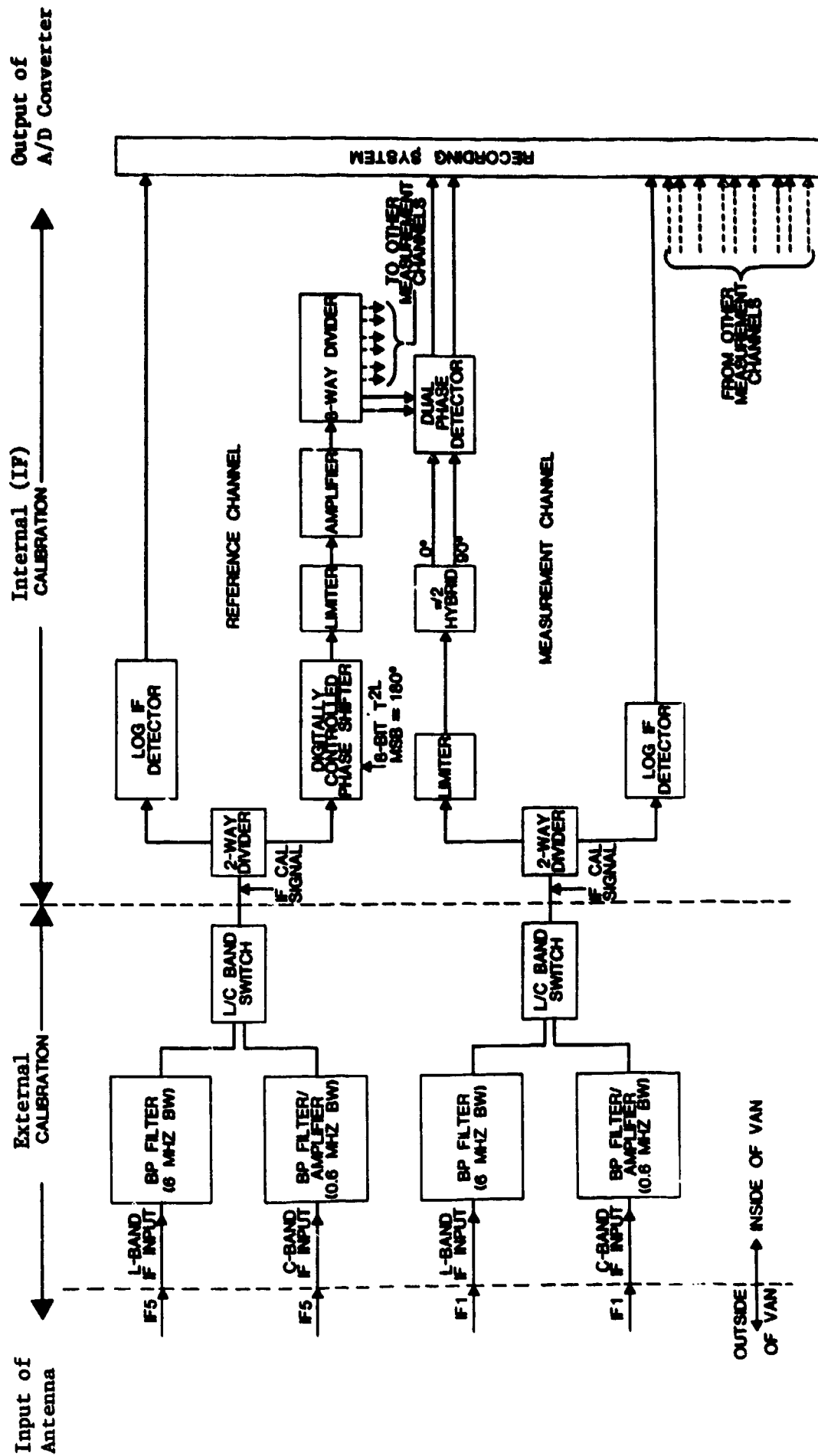
Fig. 6-3c. Block diagram of current aperture sampling system: IF receiver.

When the PALM facility was first modified for L-band terrain multipath measurements, the 5-channel receivers (both RF and IF) were kept inside the electronic van and (eleven) 50-foot RG-214/U coaxial cables were used to connect the dipole antennas to the RF front ends. The electrical path length of the RG-214/U cable was found to be very temperature sensitive at L-band. Subsequently, when the azimuth array was added, several steps were taken to minimize the system temperature sensitivity due to the RF cables.

First, the RF front ends were moved outside the van and installed behind the antenna elements. This significantly shortened the required length of RF cables from the original 50 to 10 feet. Also, this allowed the long cable connection needed between the output of the RF front end to the IF receiver inside the van to operate at a 60-MHz IF frequency, where the temperature sensitivity of RG-214/U cable would have much less effect on the channel differential stability. Second, temperature-stable semi-rigid cables and flexible cables were used for all RF connections between the antenna elements and their RF front end inputs. These cables were cut to be of equal length and to be as short as possible (10 feet of semi-rigid cable for the elevation antenna and 12 feet of cable for the azimuth antenna). With these cables, measurements indicate an electrical phase stability of better than 1° for a 40°F temperature change.

Installation of the RF front ends at the antenna array necessitated the use of long cables to bring the local oscillator (LO) signal from the van to the RF front ends. To have differentially stable LO signals for f.   ront ends, the LO signal is brought out from the van in a single RG-214,   nd split five ways at the antenna array. Again, temperature-stable cables of equal length were used to carry the distributed LO signa RF front ends at the elevation array and an equally stable cable wa the connection to the RF front end at the azimuth array.

The 6-MHz-bandwidth IF BP filters are matched in gain and phas respect to the reference receiver IF filter to within ± 0.2 dB and ± respectively. Measurements indicate a differential electrical phase stabi for the IF filters and amplifier of better than ± 1° over a 24-hour period.

6-9

b. C-band Subsystem

This subsystem uses a combination of new components and some components from the 12-foot (ITT Gilfillan) C-band Doppler elevation array developed for the FAA MLS phase II field tests [84]. The new components, including 5-channel RF front ends, 29-horn antennas, and 0.6-MHz-bandwidth IF BP filters, had tight specifications with respect to the channel matching and temperature characteristics. The components obtained from the ITT array, including RF cables, antenna switches and array housing, were measured to verify good channel matching.

The 29-horn antennas are essentially identical. The isolation between any two neighboring horns is better than 40 dB. The RF cables, connecting between the antennas to the antenna switch, are equal-length short (2 feet), and are temperature stable. For each single-pole 10-throw (SP10T) antenna switch from the ITT array, the seven most similar input ports were chosen for each RF channel. The output of the antenna switch was directly connected to the input of the RF front end with no RF cable involved. The LO signal is generated at the antenna array and distributed to all five RF front ends through the RF cables of equal length. The entire C-band subsystem, except the IF BP filters, was packaged inside the array housing, as can be seen in Fig. 6-3, with interior temperature and humidity regulation to further enhance the temperature tracking among antenna/RF receiver channels.

Again, as in the L-band case, the IF bandpass (BP) filters are matched in gain and phase with respect to the reference channel IF filter to within ±0.1 dB and ±1°. These IF filters were installed inside the van in the equipment compartment containing the 5-channel receiver. The electrical stability of these filters with time is similar to that of the L band filters.

c. IF Subsystem

This subsystem is essentially unchanged from the original PALM system [42]. Since this part of the system had been proven to be fairly stable [5, 42] and is calibrated over the full dynamic range for each experiment, the only effort made to better equalizing the five IF channels was involved gain

and offset adjustments of various log-video and operational amplifiers to yield a similar amplitude gain and phase sensitivity among five channels.

2. System Calibration

System calibration consists of an internal (IF) calibration and an external calibration, as indicated in Fig. 6-2. The IF calibration is concerned with the path from the output of the L/C-band switch down to the output of the A/D converter. This path is entirely inside the equipment van and mainly consists of the IF log-video and phase detectors, the channel multiplexer, and the A/D converter. The external calibration (for either L- or C-band) calibrates the path from the antenna down to the output of the L/C-band switch. This path contains RF cable, RF switch, RF front end, IF cable, and the IF bandpass filter.

The IF amplitude calibration is accomplished by injecting a test signal at the output of the L/C-band switch and recording the A/D converter output of the log-video detector in A/D counts. The test signal is attenuated from 0 to 80 dB in 55 steps with a digital attenuator at the output of the test signal generator. The digital attenuator was calibrated on a network analyzer and the precise attenuation values (to 0.01 dB) for each attenuation setting are stored in a lookup table. Figure 6-4 shows some examples of the IF channel amplitude response curves obtained from the IF amplitude calibrations taken at various times. The IF amplitude calibration generates amplitude lookup tables to convert the recorded signal amplitude in A/D counts to the IF signal amplitude in dB. The peak quantization error for the amplitude measurement with the current 8-bit A/D converter is about 0.2 dB.

The IF phase calibration is done by feeding the IF test signal to the output of the L/C-band switch and recording the phase detector output in A/D counts in terms of the in-phase (I) and the quadrature-phase (Q) components. The channel relative phase is varied from 0° to 360° in 32 steps by means of a digital phase shifter in the reference IF channel. In the IF phase calibration, channel #5 is used as a phase reference channel. The digital phase shifter was calibrated on a network analyzer and precise phase values (to

6-11

(a) Premission IF amplitude calibration
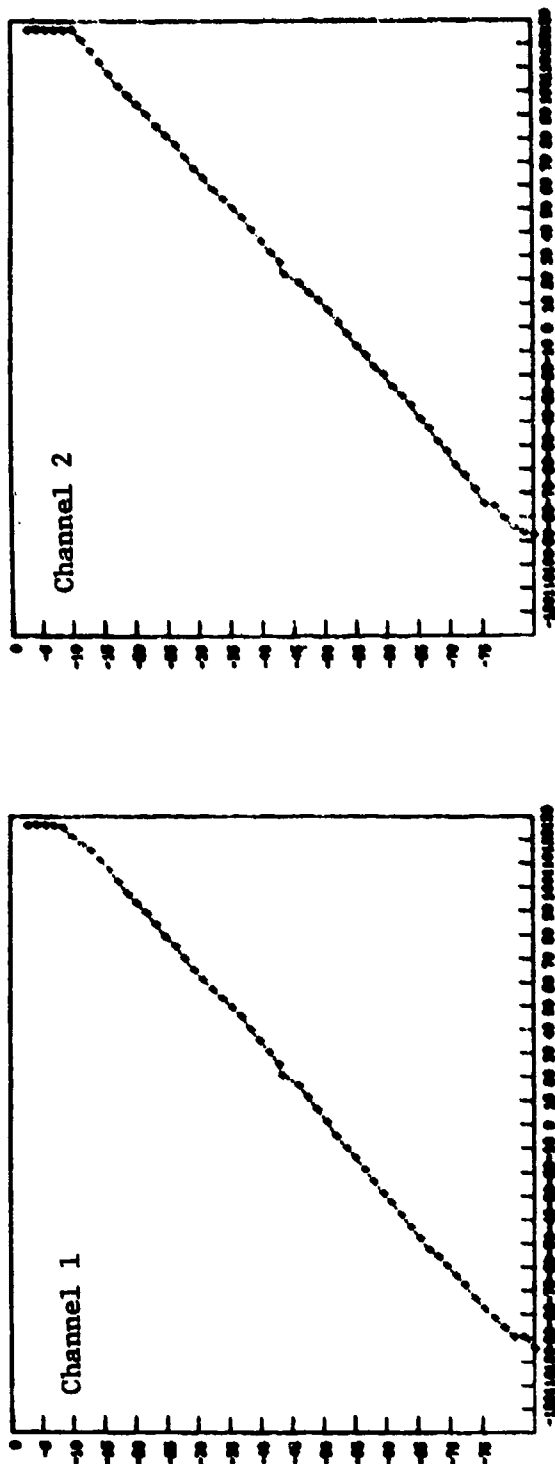
(b) Postmission IF amplitude calibration
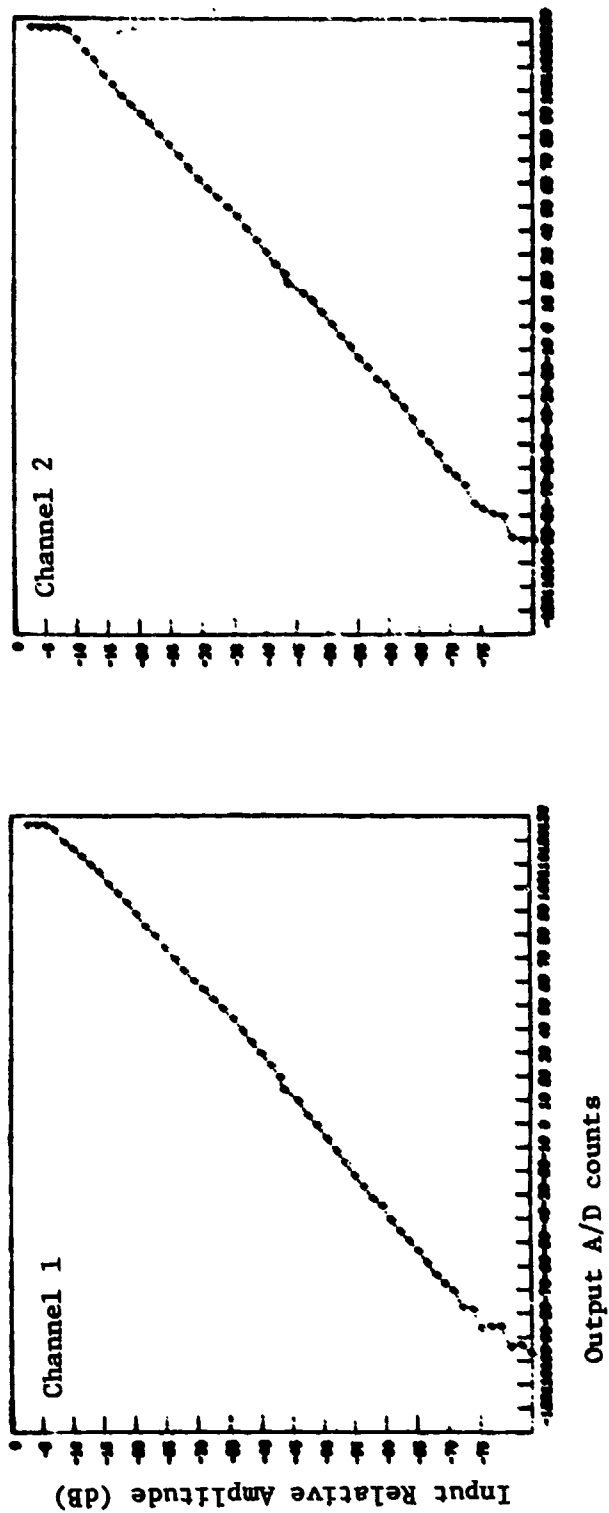
Fig. 6-4. IF channel amplitude response.

0.05°) for each digital phase setting are stored in a lookup table. Examples of the IF channel phase response curves, together with the measured I and Q values, are shown in Fig. 6-5. The saturation that appears in Fig. 6-5 on both I and Q curves was intentional so as to obtain higher A/D quantization accuracy with the same 8-bit A/D converter. The IF phase calibration yields phase lookup tables to convert the measured I and Q values in A/D counts to the relative phase in degrees. The peak phase quantization error is about 0.3°.

For a given field measurement mission, which normally lasted from 4 to 6 hours, at least two IF calibrations (premission and postmission) were made to determine the short-term IF subsystem stability. As can be seen in Figs. 6-4 and 6-5, both amplitude and phase responses remain very much the same between the premission and postmission calibrations. Similar short-term IF subsystem stability was achieved on the other experiments; although there were significant differences in some cases between the calibrations taken at different dates.

The calibration of the system from the input to the antenna to the output of L/C-band switch, i.e., the external calibration, was accomplished by the array collimation experiment shown in Fig. 6-6 at the Lincoln Laboratory antenna test range. This antenna test range was graded to have a flatness of better than ±1" over a 200-foot by 2000-foot area. The measured cross-range field strength variation[*] at the receiving end was well below 0.5 dB for either our L- or C-band arrays. At the antenna test range, the antenna array (either L- or C-band) was arranged in such a way that a common far-field source would uniformly illuminate the entire array at the same time, i.e., there would be a single plane wave normal incident upon the array. Thus, any

---

[*]For L-band, the transmitting source was an 8-foot dish at 30-foot height and the L-band elevation array (30 foot long) was laid horizontally at a 3.5-foot height above the range terrain. For C-band, the transmitting source was a 6-foot dish at 10-foot height and the C-band array (12 foot long) was laid horizontally at a 5.5-foot height.

(a) Premission IF phase calibration
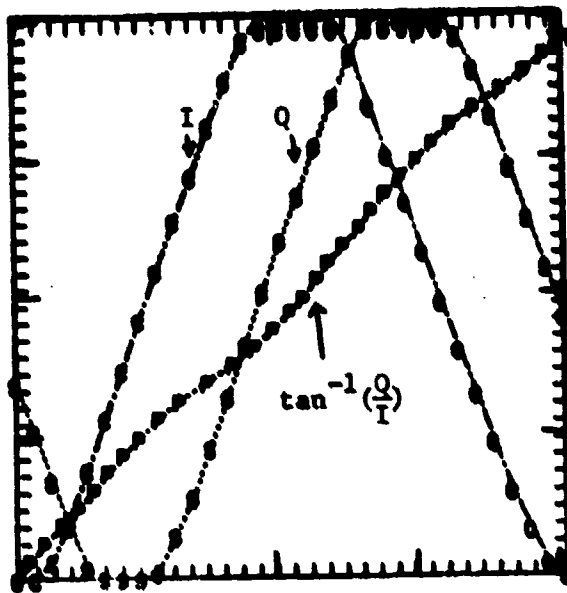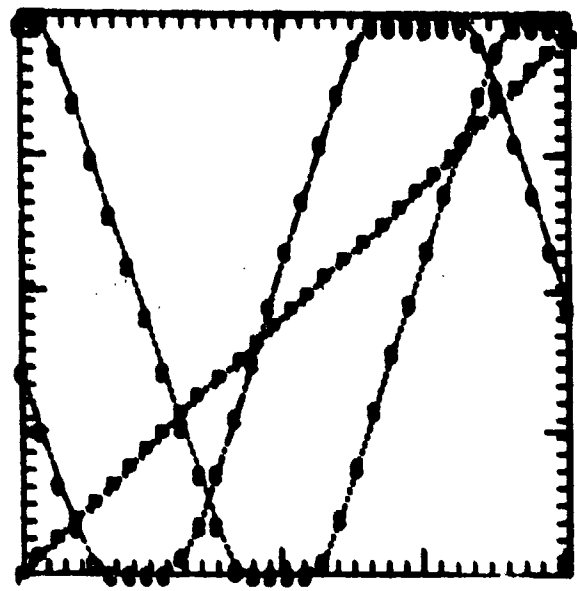
channel 1

channel 2

(b) Postmission IF phase calibration
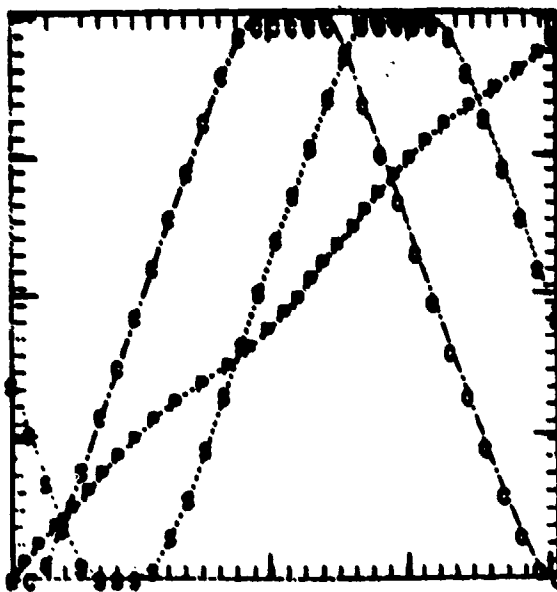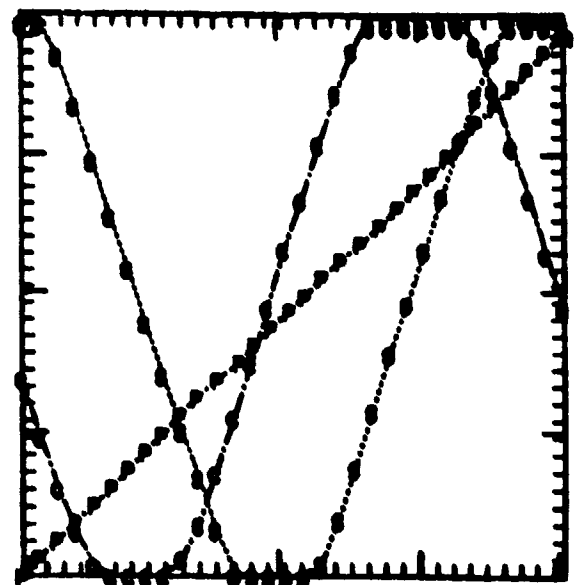
channel 1

channel 2

Fig. 6-5. IF channel phase response.
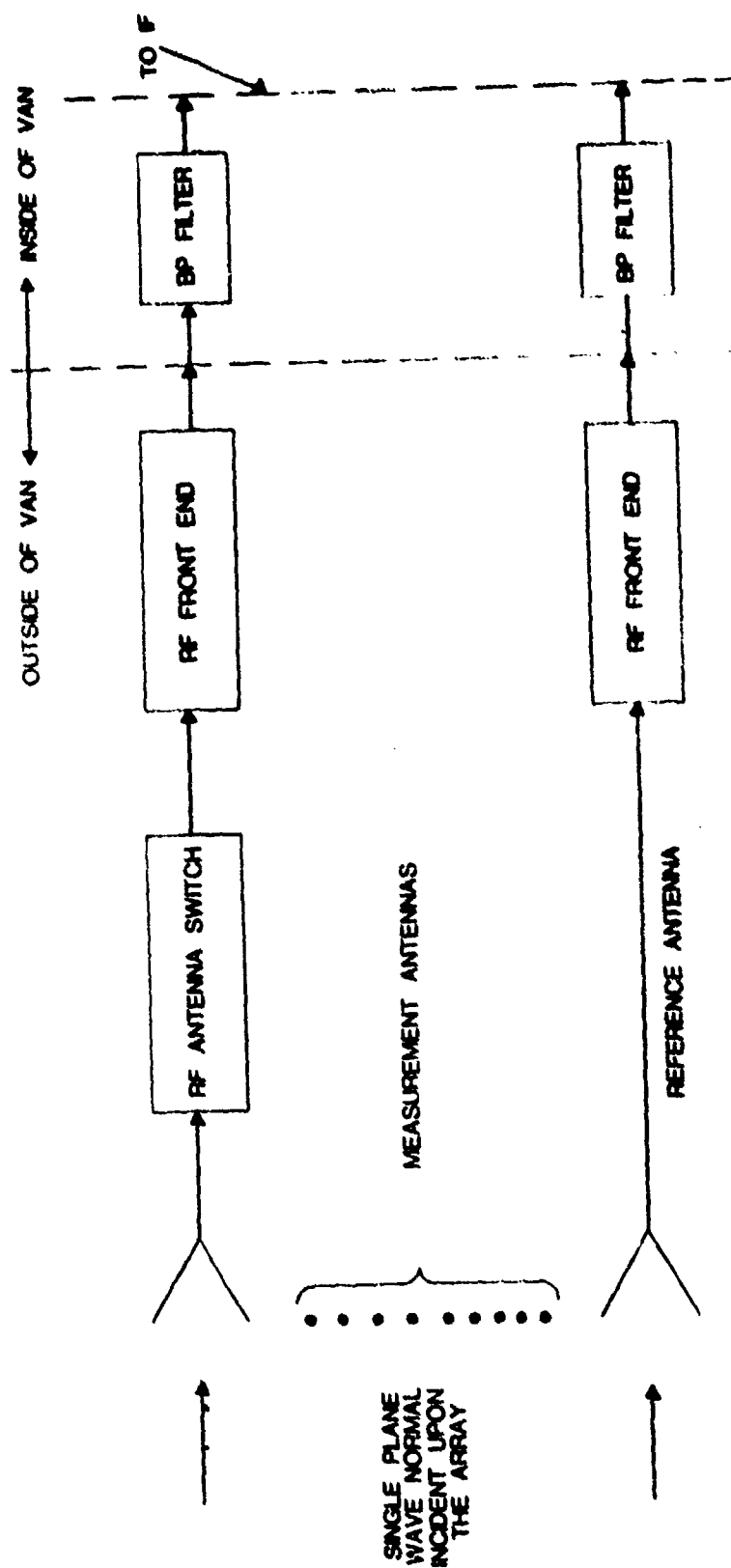
Fig. 6-6. External calibration by array collimation at antenna test range.

differences in the measured amplitude and phase values for the various antenna/receiver channels are due to amplitude and phase biases.

External calibration by array collimation (as accomplished at the antenna test range) could not be done at field measurement sites due to difficulties in achieving uniform plane wave illumination of the array in the presence of inhomogeneous terrain. The compromise used to date is to:

> (1) make an external calibration by array collimation at the antenna test range in several different times of year, e.g., summer, fall and winter. This permits one to apply appropriate external calibration data to process various field measurement data taken at different times of year, and

> (2) to compensate for drifts in the front-end subsystem by making an RF calibration as indicated in Fig. 6-3, during array collimation at the antenna test range and prior to field experiments.

The RF calibration shown in Fig. 6-3 is accomplished by injecting the RF test signal immediately behind the antenna element and recording the measured phase and amplitude for each antenna/receiver channel. In the current system, the C-band RF test signal is injected behind each of the 29-horn antennas by a 40-dB slotted-waveguide directional coupler through a common waveguide, as shown in Fig. 6-7. The directivity of the coupler is better than 20 dB. With this 20 dB directivity plus the minimum of 40 dB isolation between two neighboring horns, the injected test signal is mainly coupled back toward the output of the RF front end for the desired RF calibration. One problem encountered in the C-band RF calibration was that the RF calibration data are sometimes inconsistent with the corresponding external calibration data by array collimation. We suspect that the problem is caused by an unstable RF test signal due to the use of a waveguide feeding point which is less than one wavelength from the C-band reference channel horn. This distance probably is too short to ensure the sufficient attenuation of the higher order and/or non-propagating waveguide modes. Consequently, small temperature and humidity changes cause large changes in these higher modes which in turn can cause the test signal variations at the reference channel and those antenna channels near the feeding point.

113764-R

40 dB DIRECTIONAL COUPLERS

CALIBRATION INPUT

CLOSELY SPACED HORN ASSEMBLY (Lowest section)

RF DECOUPLING CHOKE
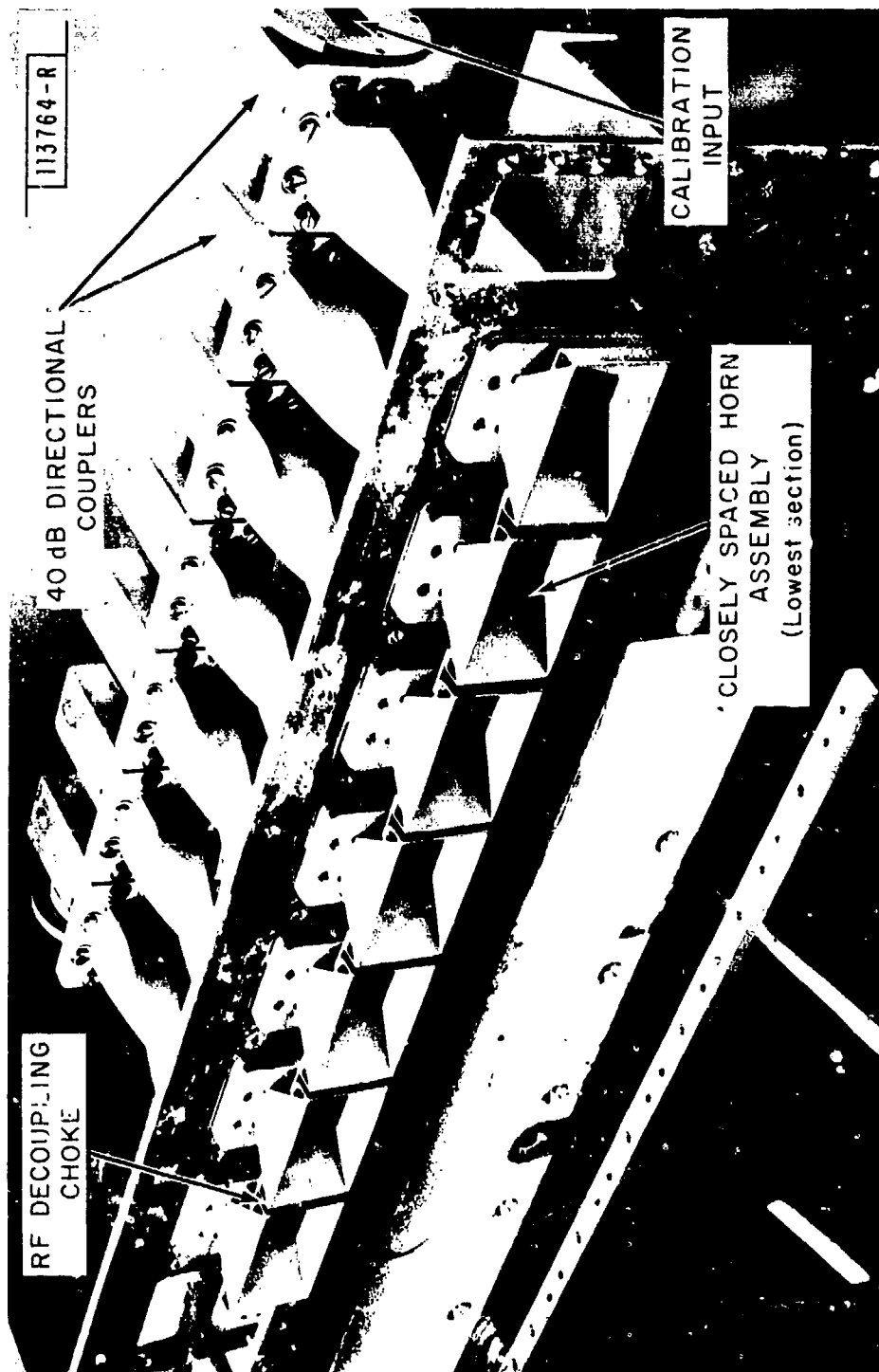
Fig. 6-7. Coupling arrangement of C-band RF test signal.

6-17

The L-band RF test signal is injected through a coaxial directional coupler to only one of the four antenna channels in the same RF front end. The choice of one antenna channel over all four antenna channels for the L-band RF test signal injection was made based on the following arguments:

(1) It is much simpler to distribute the RF test signal in the one antenna channel case than in the four antenna channel case (one 5-way power divider, five 10-foot RF cables, and five directional couplers versus one 17-way power divider, 17 15-foot RF cables, and 17 directional couplers). In particular, the temperature tracking among test signal paths should be better for the one antenna channel case.

(2) Since the short RF cables in four antenna channels which share the same RF front end are of equal length and very temperature stable, RF calibration for one antenna channel should represent the situation for the other three antenna channels using the same RF front end and

(3) RF cable electrical length variations with temperature should be much smaller at L- than C-band.

By comparing the RF calibration values taken at any given measurement mission with those taken at the external calibration by array collimation at the antenna test range, any change in the channel phase and amplitude biases obtained from the external calibration can be compensated.

Table 6-1 shows two sets of external calibration values taken about six hours apart for the C-band elevation antenna system. Also shown in this table are the corresponding RF calibration values. Table 6-2 gives the similar results for the L-band antenna system which has 11 elevation and 6 azimuth antenna/receiver channels. Additional calibration data are given in Appendix C. Examination of the external array collimation data and the corresponding RF calibration data obtained during the past year showed that both the C- and L-band elevation antenna systems had been fairly stable in the short term (4-6 hours), while the L-band azimuth antenna system showed notable drift in the channel phase bias which often could not be explained by the corresponding RF calibration values. This larger channel phase bias variation observed in the L-band azimuth array channels may be is due to the antenna elements. Contrary to the simple horn antenna used in the C-band elevation antenna system and the

6-18

TABLE 6-1

EXTERNAL AND RF CALIBRATION VALUES:
C-BAND ANTENNA SYSTEM

Date: 8/15/80

(a) External calibration values
AMP1/AMPR and PH1/PHR: relative amplitude and phase values from two
external calibrations taken 6 hours apart

ANT	AMP1	AMPR	AMP1-AMPR	PH1	PHR	PH1-PHR
2	-0.15	-0.30	0.15	865.54	868.99	-3.45
3	-0.81	-0.98	0.17	183.44	130.37	-3.93
4	-1.83	-2.10	0.27	165.77	166.11	-3.34
5	-1.83	-1.39	0.16	244.71	248.12	-3.41
6	-0.30	-0.30	0.00	845.88	849.38	-3.48
7	-0.81	-0.98	0.17	290.74	293.70	-2.96
8	-0.93	-1.05	0.18	270.78	273.32	-2.54
9	0.04	-0.30	0.34	277.78	278.59	-1.81
10	-0.63	-3.84	0.21	297.13	298.71	-1.58
11	0.04	-0.27	0.31	286.47	288.13	-1.66
12	0.42	-0.03	0.45	291.45	292.99	-1.54
13	1.57	1.23	0.34	304.71	305.73	-1.02
14	0.42	0.12	0.30	208.07	283.66	-1.59
15	-0.84	-0.93	0.09	863.57	865.34	-1.77
16	1.83	1.23	0.60	874.56	875.36	-0.80
17	1.27	0.90	0.37	301.51	302.47	-0.96
18	0.83	0.51	0.32	297.44	298.09	-0.65
19	0.38	0.00	0.38	312.13	312.96	-0.82
20	0.80	0.51	0.28	300.56	300.97	-0.41
21	0.83	0.48	0.35	325.81	325.98	-0.17
22	0.83	0.51	0.32	305.09	305.64	-0.55
23	3.03	2.82	0.21	272.40	273.77	-1.37
24	2.82	2.76	0.06	276.88	278.42	-1.54
25	1.83	1.56	0.27	855.45	857.52	-2.07
26	1.50	1.26	0.24	863.74	865.39	-1.65
27	2.49	2.55	-0.06	303.80	305.82	-1.92
28	2.06	1.89	0.17	249.89	298.59	-1.30
29	1.17	0.90	0.27	836.66	838.11	-1.45
30	0.0	0.0	0.0	0.00	0.00	0.00

(b) RF calibration values
AMP1/AMPR and PH1/PHR: relative amplitude and phase values from two
RF calibrations taken 6 hours apart

ANT	AMP1	AMPR	AMP1-AMPR	PH1	PHR	PH1-PHR
2	0.71	0.04	0.67	861.83	863.60	-1.77
3	-0.53	-1.81	0.60	44.18	44.36	-0.44
4	-0.88	-2.74	0.88	846.88	845.18	-0.16
5	-0.88	-1.69	0.80	46.88	47.18	-0.57
6	-0.88	-0.97	0.88	894.73	895.18	-0.39
7	-1.13	-1.93	0.80	94.96	96.49	-1.58
8	-1.10	-1.87	0.77	858.49	854.56	-0.07
9	1.36	0.50	0.86	79.88	81.98	-0.15
10	0.34	-0.35	0.69	878.84	880.83	-1.99
11	-0.38	-1.10	0.72	364.99	367.81	-2.28
12	0.34	-0.42	0.76	92.80	93.78	-1.58
13	1.44	0.58	0.86	195.86	197.10	-1.84
14	-0.15	-1.02	0.87	884.46	886.12	-1.66
15	-1.04	-1.96	0.92	338.66	334.79	-0.13
16	1.32	0.68	0.70	74.80	74.80	0.01
17	0.61	-0.19	0.80	193.28	193.46	-0.18
18	0.55	-0.19	0.74	878.88	880.88	-1.45
19	-0.56	-1.59	0.98	19.88	81.88	-1.73
20	0.79	0.12	0.67	108.34	108.38	0.04
21	0.91	0.19	0.71	816.43	817.88	-1.13
22	0.79	0.00	0.78	884.87	884.87	-0.07
23	2.69	2.29	0.40	338.94	339.88	-0.88
24	2.89	2.24	0.65	90.83	90.84	-0.60
25	2.69	2.24	0.45	150.38	152.81	-0.45
26	1.48	0.95	0.68	848.14	851.41	-3.87
27	2.64	2.19	0.45	356.88	358.86	-1.98
28	2.21	1.44	0.77	88.80	89.43	-0.17
29	1.24	0.47	0.78	131.88	134.04	-0.35
30	0.0	0.0	0.0	0.00	0.00	0.00

TABLE 6-2

EXTERNAL AND RF CALIBRATION VALUES:
L-BAND ANTENNA SYSTEM

Date: 7/23/80

(a) External calibration values
AMP1/AMPR and PH1/PHR: relative amplitudes and phase values from
two external calibrations taken 6 hours apart

	ANT	AMP1	AMPR	AMP1-AMPR	PH1	PHR	PH1-PHR
elevation array	9	-3.24	-3.66	0.62	391.61	391.76	-0.15
	8	-3.65	-3.64	0.43	366.56	367.26	-0.49
	7	-3.00	-3.71	0.65	341.10	341.61	-0.49
	6	-1.80	-3.36	0.40	365.31	366.43	-1.19
	5	-0.10	-0.37	0.47	338.63	340.76	-1.19
	4	-0.83	-0.61	0.37	193.66	194.61	-1.20
	3.5	-0.62	-1.18	0.82	811.66	813.36	-1.66
	3	-0.60	-0.10	0.41	346.69	348.63	-3.14
	2.5	-1.41	-1.56	0.16	146.60	147.73	-8.11
	2	-1.66	-1.39	0.31	803.19	805.07	-1.86
	1	0.00	0.00	0.00	0.00	0.00	0.00
azimuth array	6	-13.64	-13.56	0.96	78.90	80.72	-7.82
	5	-13.86	-14.16	0.90	181.30	120.84	-8.56
	4	-18.33	-18.37	0.04	825.76	833.91	-7.13
	3	-16.49	-17.04	0.54	305.66	317.18	-18.10
	2	-10.27	-10.67	0.40	11.83	18.68	-7.60
	1	-11.41	-11.97	0.55	873.96	881.66	-7.96

(b) RF calibration values
AMP1/AMPR and PH1/PHR: relative amplitude and phase values from two
RF calibrations taken 6 hours apart

	ANT.	AMP1	AMPR	AMP1-AMPR	PH1	PHR	PH1-PHR
elevation array	9	-2.89	-4.06	1.16	313.37	313.30	0.07
	8	-2.89	-4.06	1.16	313.37	313.30	0.07
	7	-2.89	-4.06	1.16	313.37	313.30	0.07
	6	-2.82	-4.06	1.16	313.37	313.30	0.07
	5	-0.38	-1.50	1.13	868.84	872.90	-3.86
	4	-0.38	-1.50	1.13	868.84	871.90	-3.66
	3.5	-0.38	-1.50	1.13	868.84	871.90	-3.66
	3	-0.38	-1.50	1.13	868.84	871.90	-3.66
	2.5	-1.50	-3.19	1.69	810.81	813.00	-8.19
	2	-1.50	-3.19	1.69	810.81	813.00	-8.19
	1	0.00	0.00	0.00	0.00	0.00	0.00
azimuth array	6	-9.85	-10.93	1.88	116.81	105.88	-9.87
	5	-9.85	-10.93	1.88	116.81	105.88	-9.87
	4	-9.85	-10.93	1.88	116.81	105.88	-9.87
	3	-9.85	-10.83	1.88	116.81	105.88	-9.87
	2	-8.99	-10.46	1.47	303.50	305.75	-8.85
	1	-8.99	-10.46	1.47	303.50	305.75	-8.85

6-20

conventional dipole antenna used in the L-band elevation antenna system, the L-band azimuth elements are printed-circuit dipole arrays whose more complicated feed structure apparently makes it more sensitive to temperature and humidity changes.[*]

Table 6-3 summarizes the functions of the above-discussed system calibration procedure in removing the various channel tracking error sources. Figure 6-8 shows how various system calibration data are used to obtain the correct signal amplitude and phase at each antenna element, i.e., the measured amplitude and phase values which are free of channel amplitude and phase biases. Judging from the system calibration data accumulated to date and from simulation results corresponding to various field measurements with a known signal environment, the accuracy in the estimated amplitude and phase at an individual antenna element is about 1 dB and 3 - 5 degrees, respectively. These errors are believed to be primarily from the RF subsystem drifts biases which were not compensated for with the current RF calibration and/or short-duration changes in the IF subsystem.

D. Recommendations for Hardware Implementation of Advanced Sampled Aperture Systems

In this section, we consider some of the principal hardware implementation issues which arise in attempting to implement the advanced signal processing techniques discussed in the preceding sections. In Chapter V, we saw that to achieve significant improvements in performance against inbeam multipath (e.g., factors of 2-4 reduction in the region of significant errors), it is necessary to have tighter control of amplitude and phase errors than is the case with "conventional" processing. Thus, it is important to identify the major difficulties which may arise in practical implementations.

---

[*]Much of the instability of these elements is believed to arise from the printed-circuit substrate which expands in high-humidity conditions. The degree of expansion/contraction has been great enough on several occasions to cause failure of the printed-circuit lines.

6-21

TABLE 6-3

SYSTEM CALIBRATION[*]

Error Source	Removed by
IF channel bias (amplitude and pnase)	IF calibration
Antenna/RF channel bias	External Calibration by array collimation at antenna test range
RF channel drift**	RF Calibration
Antenna drift**	Not accounted for

*Not a real time calibration
**Drift refers to tne change between the time that external calibration was made and the time a given field measurement was taken.
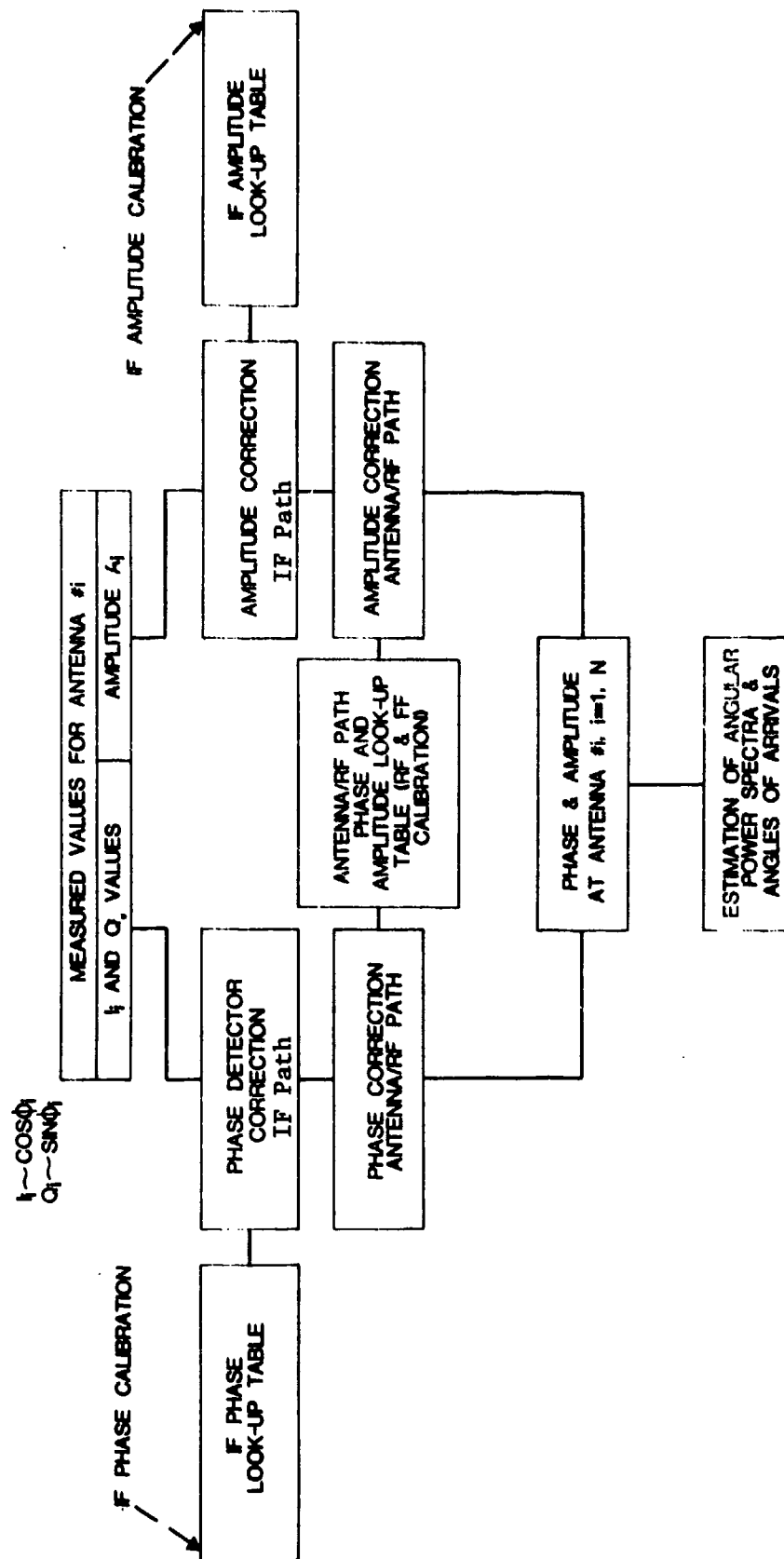
Fig. 6-8. Correction of channel amplitude and phase biases from the measured data.

1. Antenna Elements

Ideally, the individual array element patterns would be identical to within a constant gain factor. However, this is not always easy to achieve due to 1) differences between the isolated element patterns and 2) coupling between the various elements.

Differences between the isolated element patterns in the non-scanned plane* are particularly undesirable since there is no hope of compensating for them using the array data alone. The degree of similarity amongst the several printed-circuit arrays used as the L-band azimuth array element is shown in Fig. 6-9. We see that the mainlobe region is quite consistent, but that differences of 1-2 dB can occur near the horizon where the pattern is rolling off sharply. Comparable data have not been published for the slotted-waveguide arrays typically used for C-band MLS azimuth arrays; however, based on the fragmentary available data [31] we would expect to see similar variations for that implementation as well.

Variations in the isolated element characteristics within nominal scan plane should be small if these are simple elements (e.g., horns or dipoles). Elements with complicated feed structures in the scan plane (e.g., COMPACT elevation arrays [28]) will probably have a reproducibility similar to that of Fig. 6-9.

Coupling effects due to adjacent elements can be pronounced if the elements are spaced 1λ or less apart. This phenomenon is particularly noticeable in azimuth arrays with dipole elements when using vertical polarization since a coplanar orientation maximizes mutual coupling effects. The coupling per se is a problem only if the coupling environment seen by the individual elements is sufficiently different to produce element pattern differences. Figures 6-10 and 6-11 show the differences in gain and phase at azimuths of 0° and 60° between the various dipoles of a 21-element azimuth array with $\lambda/2$ and 1λ element spacings, respectively. We see the differences are most pronounced

---

*e.g., elevation plane for an azimuth array.

6-24
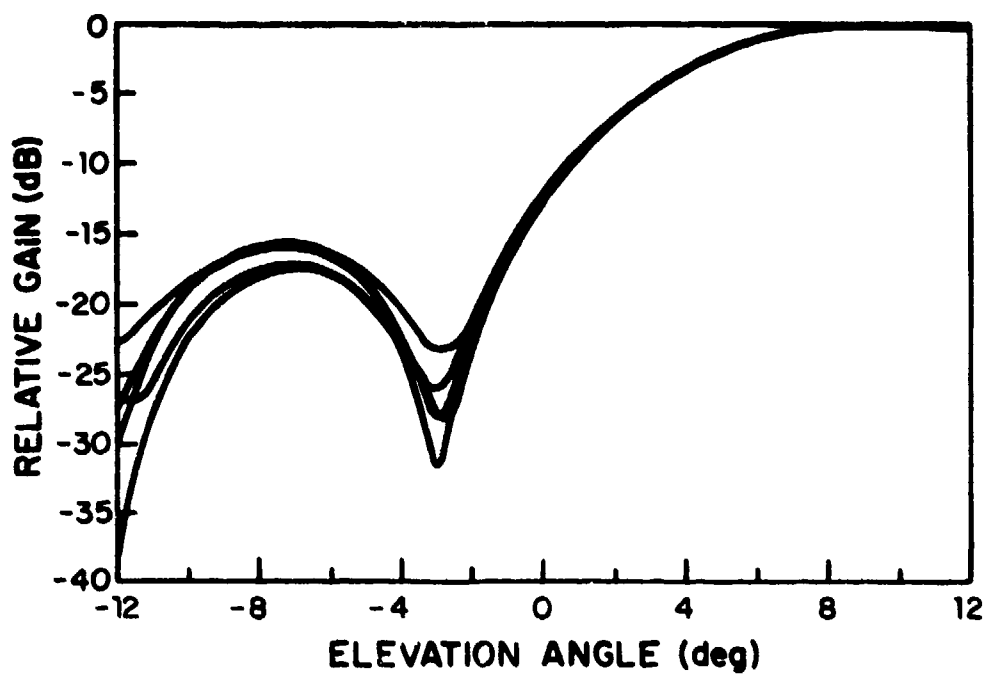
Fig. 6-9. Elevation patterns for the five PALM antennas.
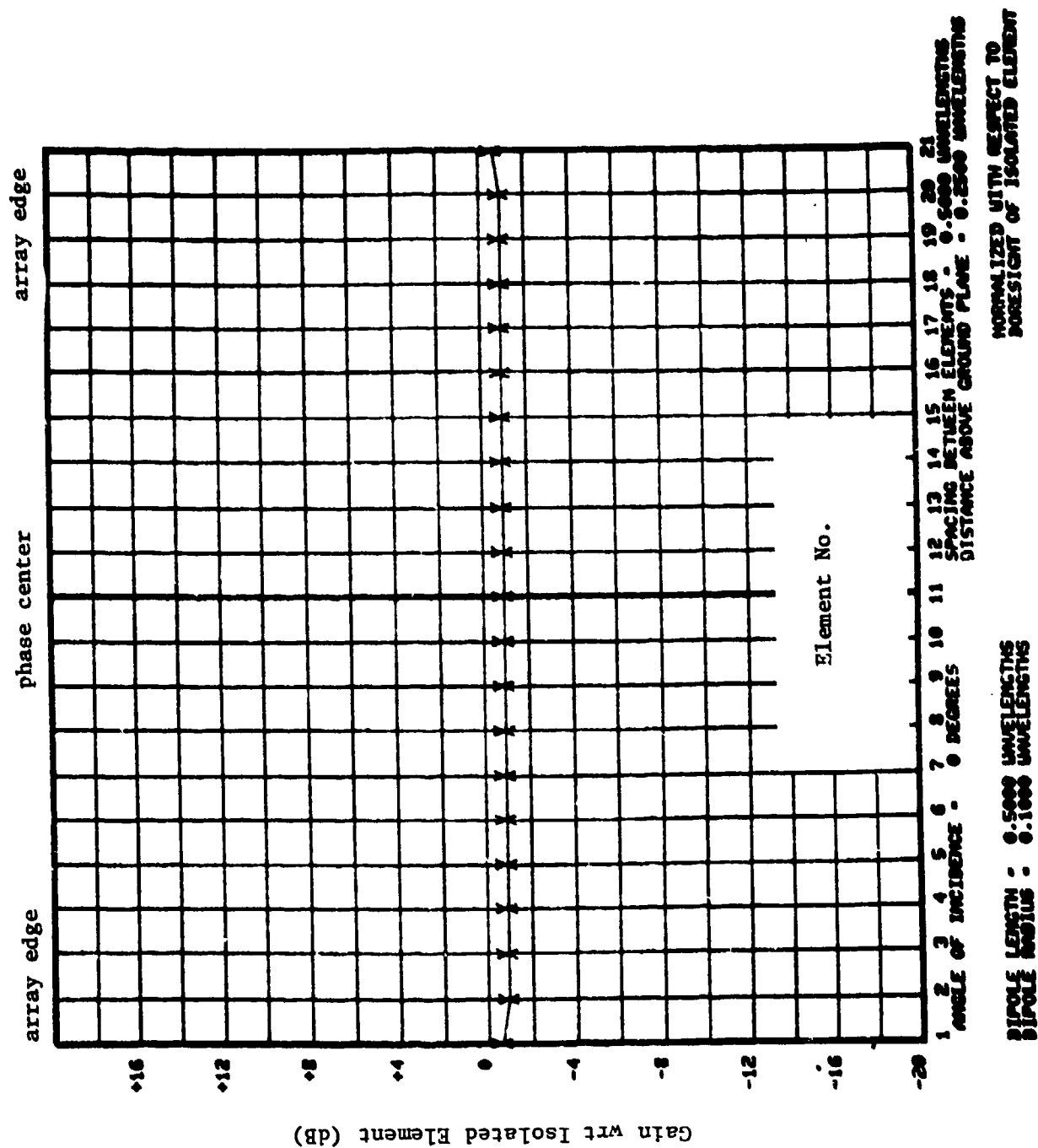
Fig. 6-10a. Gain variation due to mutual coupling on boresight for 0.5λ spacing.
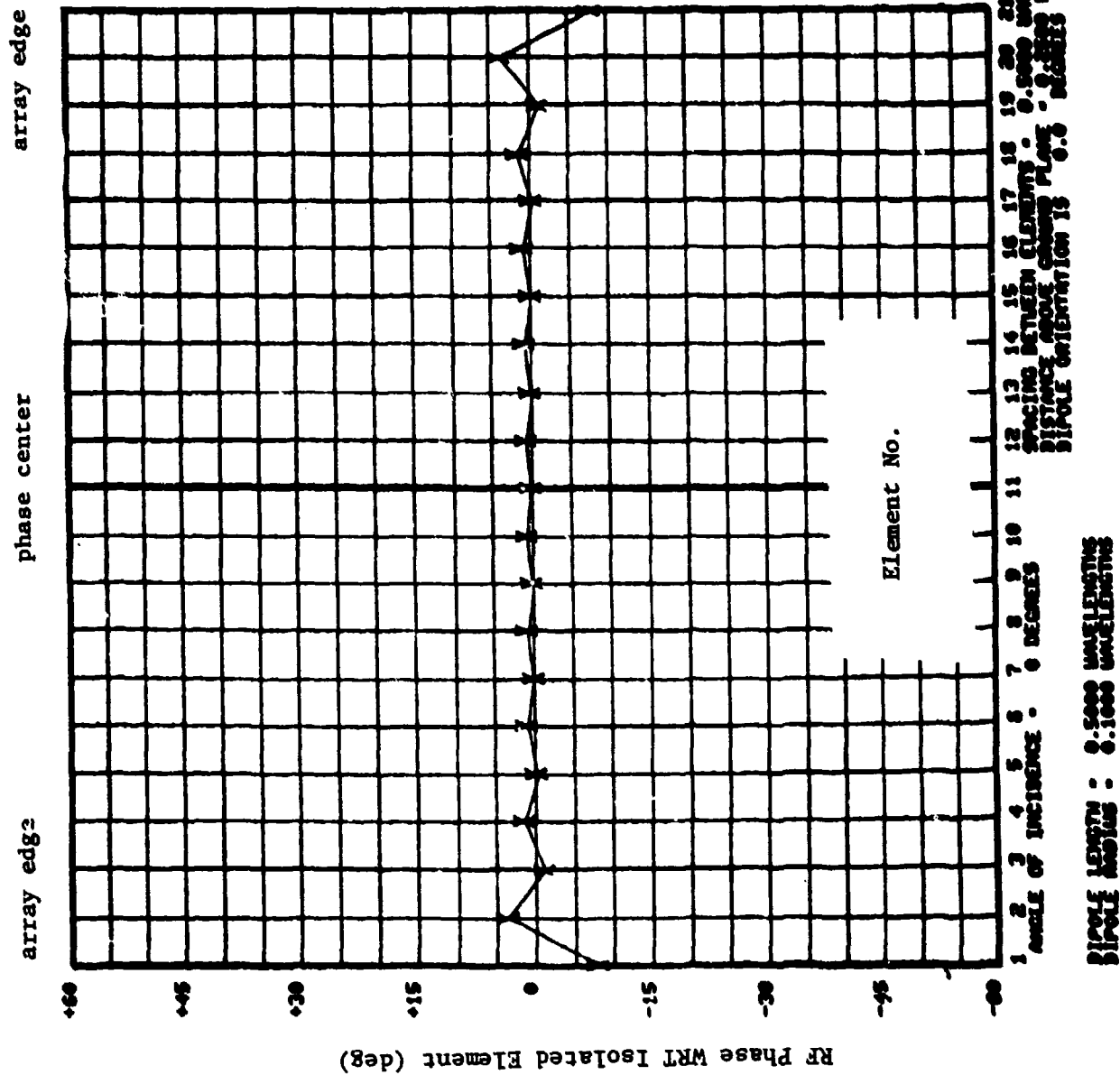
Fig. 6-10b. Phase variation due to mutual coupling for 0.5λ spacing.
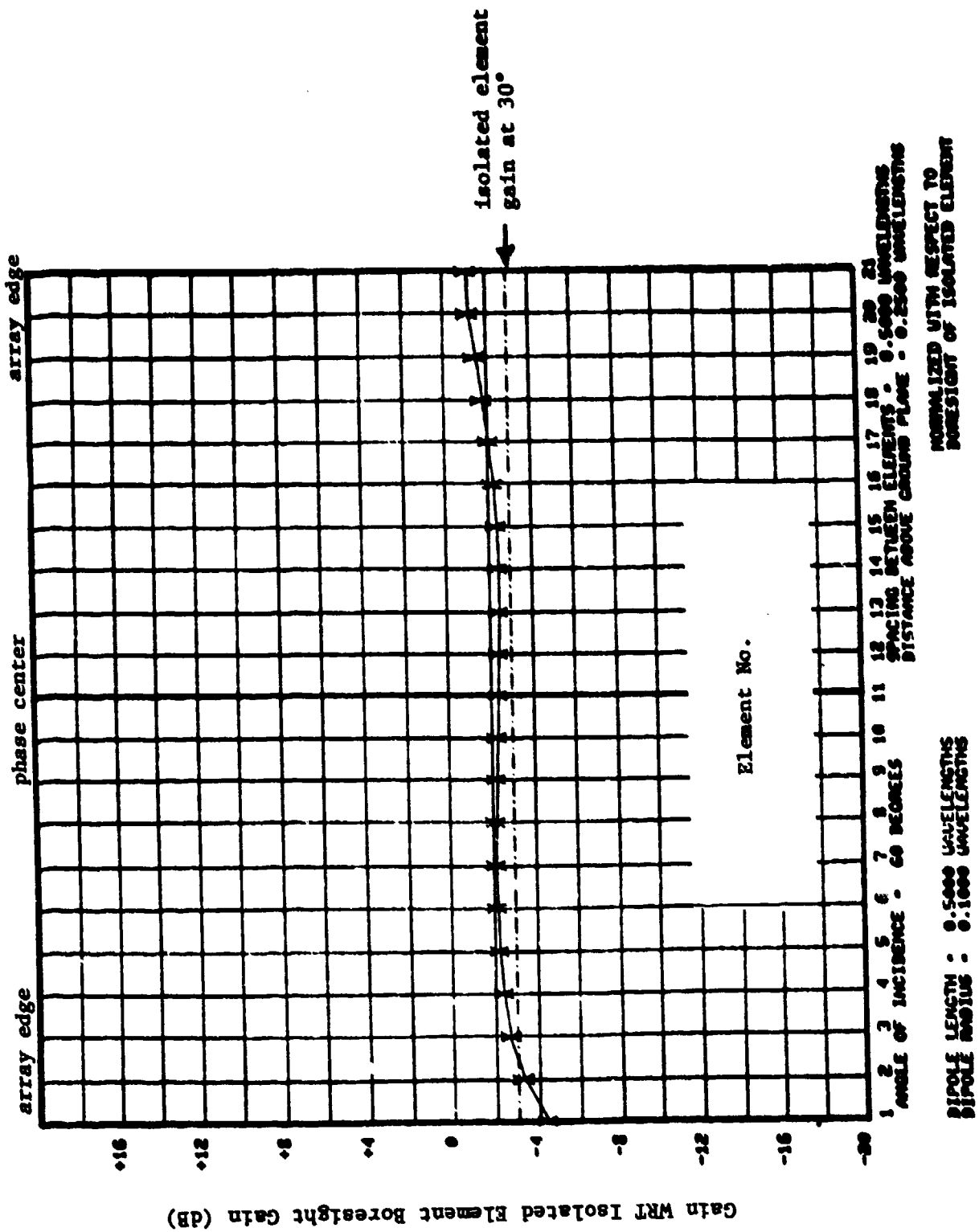
Fig. 6-10c. Gain variation due to mutual coupling at 60° off boresight for 0.5λ spacing.
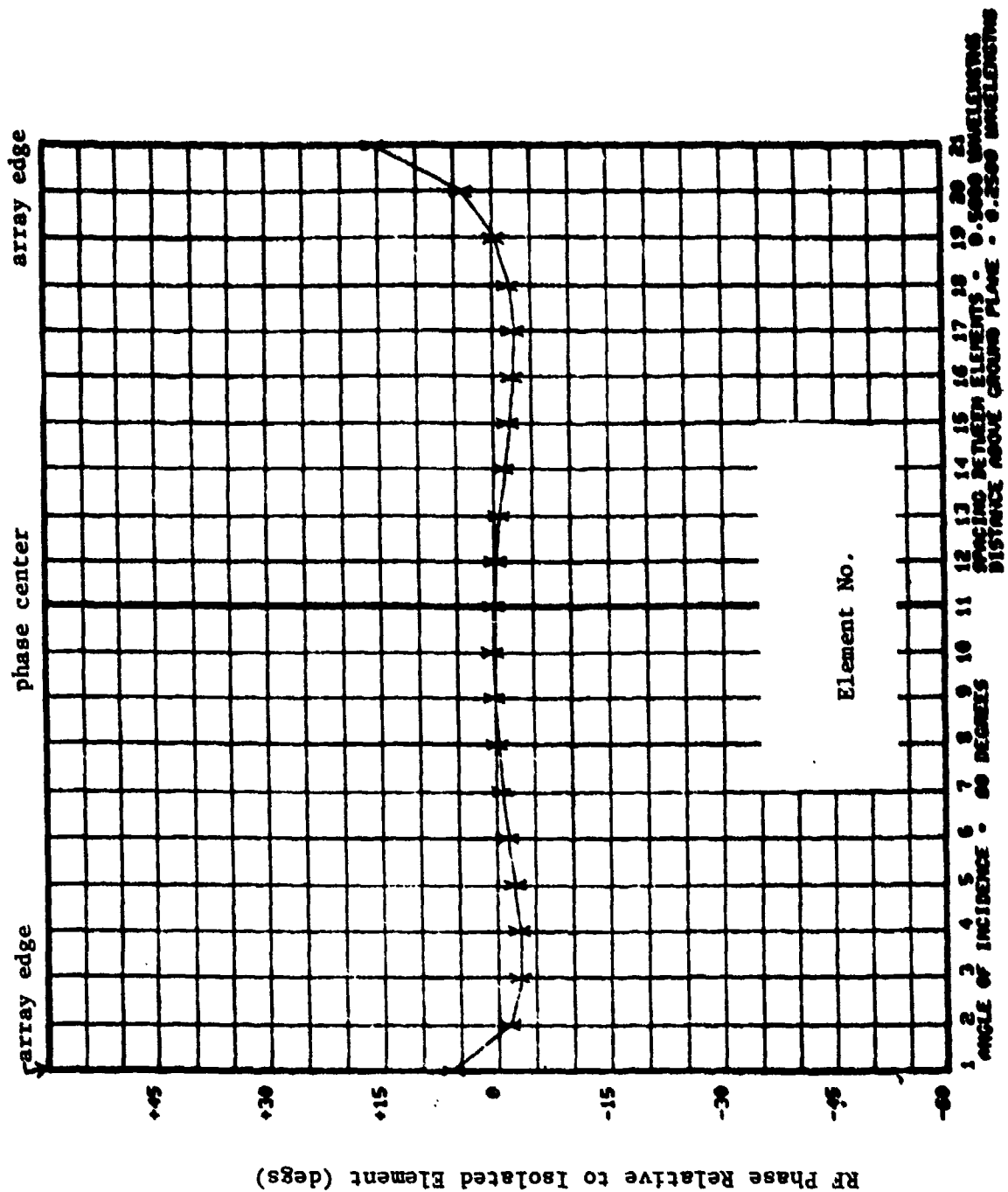
Fig. 6-10d. Phase variation due to mutual coupling at 60° off boresight for 0.5λ spacing.
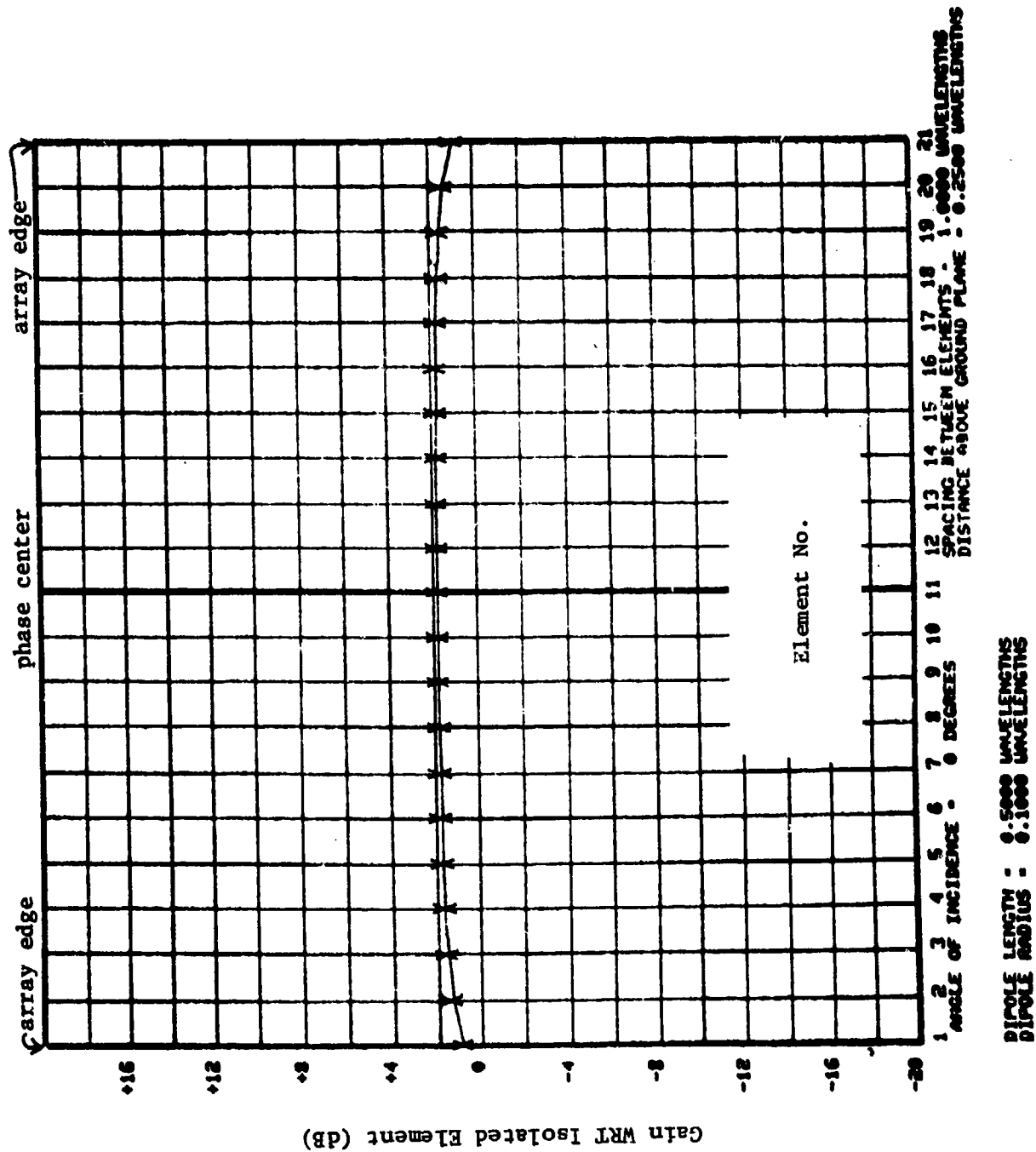
Fig. 6-11a. Gain variation due to mutual coupling on boresight for 1λ element spacing.

6-30
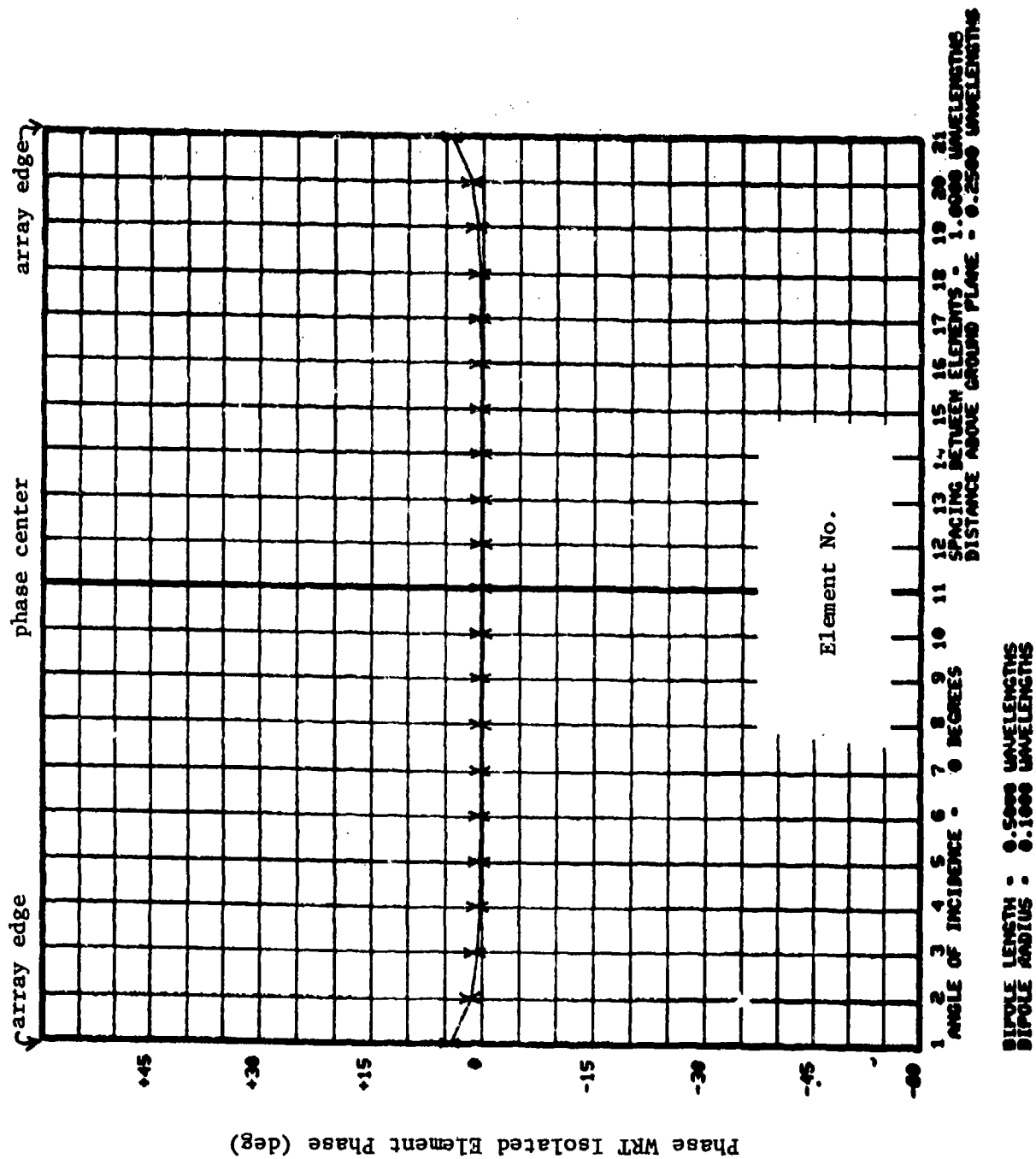
Fig. 6-11b. Phase variation due to mutual coupling on boresight for 1λ element spacing.
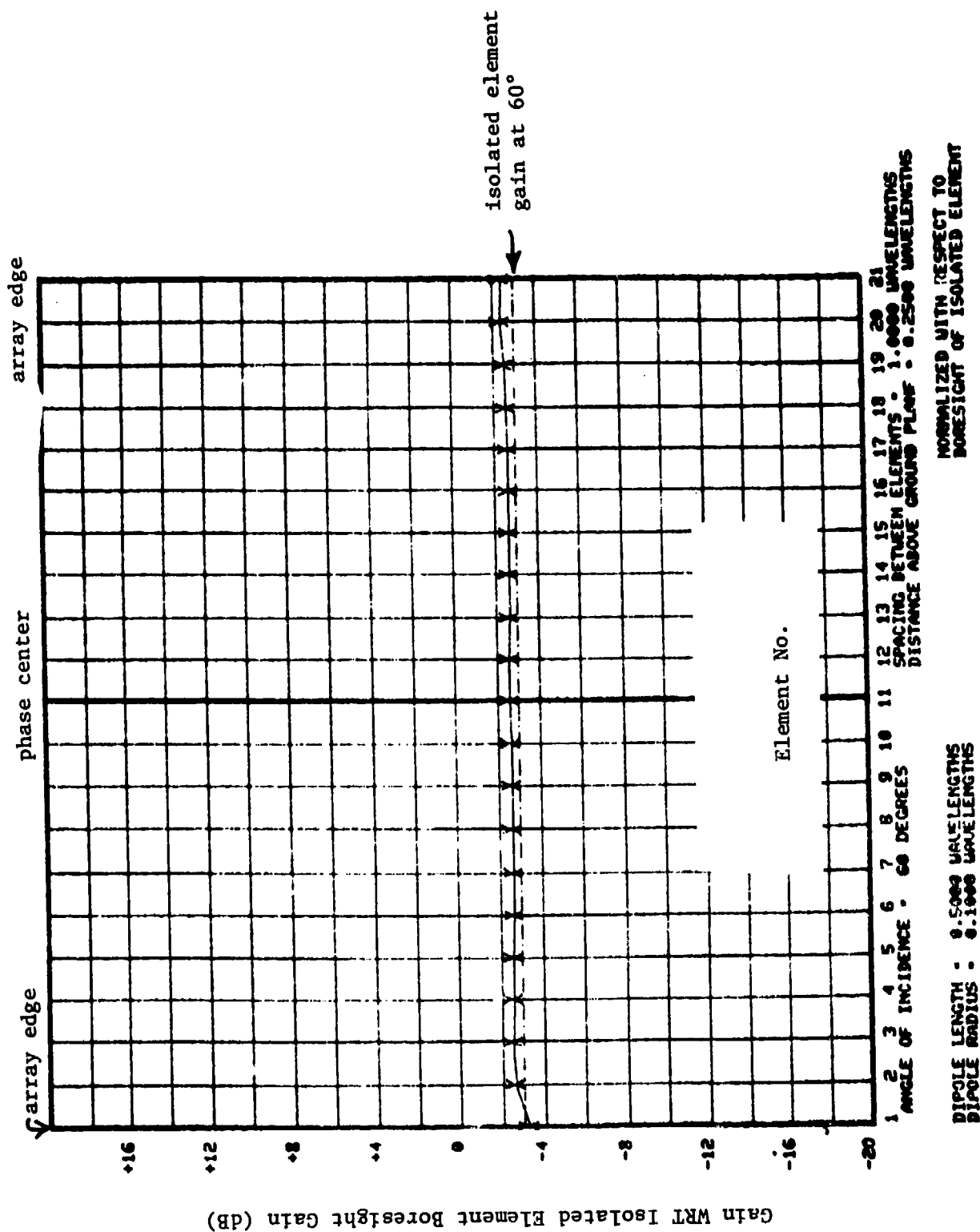
Fig. 6-11c. Gain variation due to mutual coupling at 60° for 1λ element spacing.
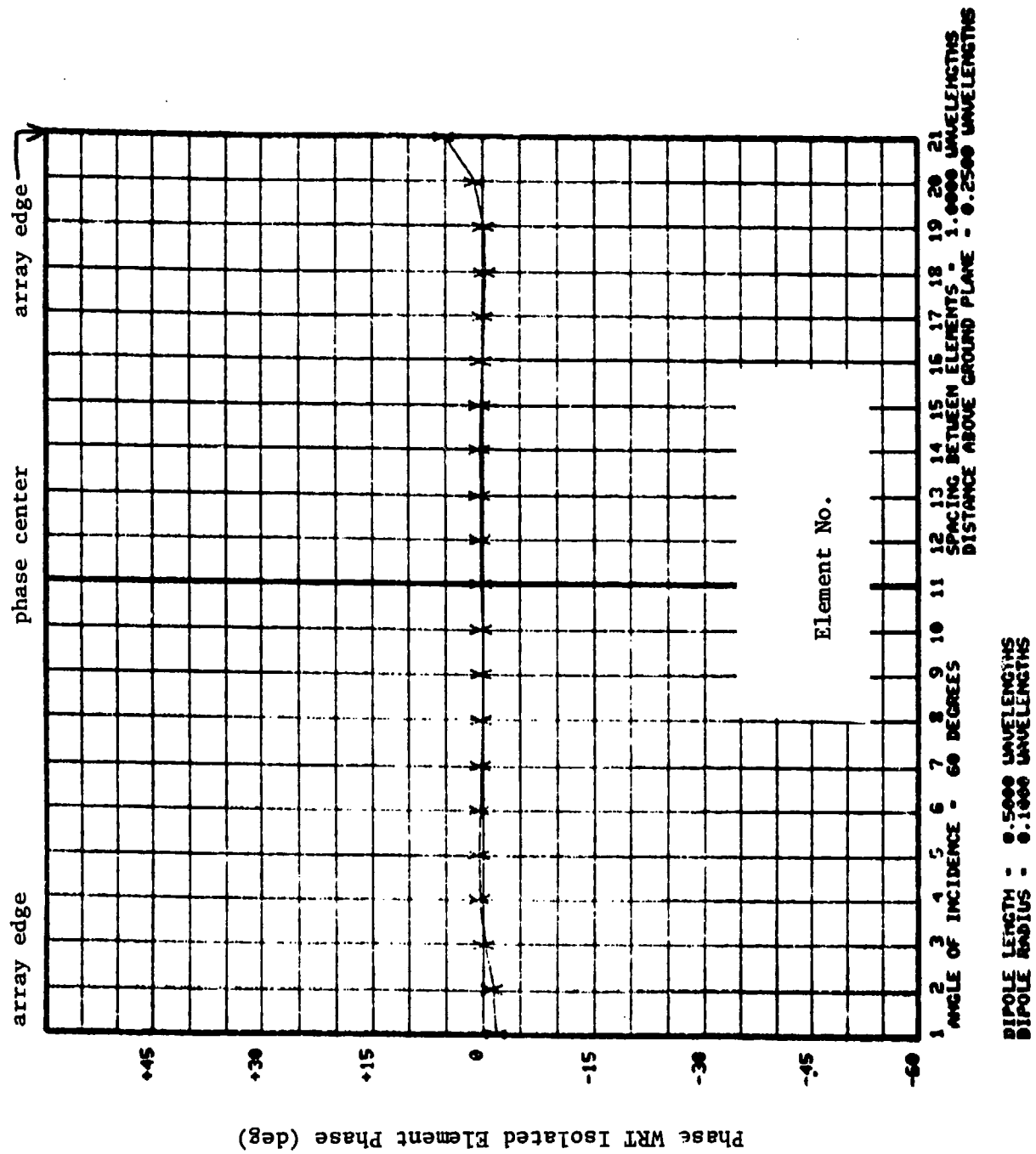
Fig. 6-11d. Phase variation due to mutual coupling at 60° for 1λ element spacing.

near the array ends which suggests using dummy elements at the array ends to mitigate the pattern differences. This use of dummy elements is easily accomplished unless it is essential that the overall array aperture be small (e.g., a few wavelengths) in which case coupling effects could seriously degrade the array performance.

2. Electrical Path Length Equalization

Variation in electrical path lengths amongst the various elements will cause phase errors in the received data. In the experimental data discussion above, we noted the need to use cabling with low temperature sensitivity. The use of front ends at the antenna elements should also aid in achieving reduced sensitivity to cable length changes since the electrical wavelength at typical IF frequencies is quite large.

3. IF Filter Characteristics

The IF filter design represents a key design issue in sampled-aperture designs since one must trade off SNR and adjacent channel interference performance (maximized by using a narrow bandwidth) against IF filter differences (which are worse near the edges of the passband). The problem is particularly troublesome if the frequency uncertainty between ground and air local oscillator frequencies is large compared to signal bandwidth. To illustrate, in the experimental system described above, it was possible to calibrate out any mid-passband differences between the L- band IF filters, but the filter gains could differ by as much as 3 dB at the passband edge. The nominal L-band signal bandwidth is on the order of 2 MHz with a nominal frequency stability of approximately 0.5 MHz. However, experimental measurements have shown that L-band transponder frequencies for the total aircraft population can vary as much as 3 MHz. Thus, to avoid measuring some transponder signals at the passband edges, it would be necessary to use an IF filter bandwidth several times larger than that required for the pulses alone.*

---

*Since the actual experimental equipment only needed to measure signal from a single transponder, we chose to use a transponder with crystal-controlled (cont'd.)

One approach to alleviating this problem would be to use a finite impulse response (FIR) IF filter design since these typically give better filter characteristic control at the band edges as well as linear phase. The FIR filter can be realized either by analog methods (SAW, CCD) or digitally, depending on the signal bandwidth.

4. System Calibration/Compensation/Monitoring

Based on our experimental system experience it appears essential that the various signal paths be monitored more or less continously to avoid phase variations. The amplitude variations seem to be somewhat slower varying and hence may not need to be monitored as frequently. The combination of IF calibration for the individual channel amplitude and phase measurement circuits together with signals injected at the antennas (i.e., rf calibration) was reasonably successful, and would be recommended for any future systems. However, there were some discrepancies between the rf calibration and the array collimation experiments which probably represent a combination of

1) changes in the antenna element patterns (in the case of the L-band array) and

2) changes in the signal injection system (in the case of the C-band array).

A standard suggestion for reducing these effects is to transmit a signal external to the array and compare the various received signals. This was considered for the experimental arrays, but discarded because it appears dubious that the propagation paths from each of the elements to the external source would have the desired stability to permit use " real time" element compensation.

6. Received Signal Digitization

At the time when the experimental aperture sampling array was first constructed, the A/D converters represented an important constraint in that

6-35

the requisite speed and accuracy (i.e., number of bits) was costly if available at all. The size and power consumption of these A/D converters also was a major detriment to actual implementation.

However, recent rapid progress in this area has lessened the A/D problems significantly. The experimental van uses an 8-bit A/D converter which can be operated at 30 MHz, gives generally good results at 20 MHz, and is quite reliable at 10 MHz. The cost of these units is approximately $850 with the A/D converter, input buffer reference voltage and power conditioning occupying a single 4.5" x 5.5" board.

If greater A/D accuracy is desired (e.g., for I, Q sampling as opposed to the log amplitude, relative phase sampling utilized in the experimental equipment), a 12-bit 5-MHz converter fitting on a 5" x 5.5" card is currently available for $3500. Alternatively, a number of 12-bit 500-kHz converters (cost approximately $269/each) could be used. The point to be made here is that the cost and size of high-accuracy A/D converters have dropped significantly in the past few years such that the A/D converter may not be a major system constraint.

7. Hardware Realization of Mathematical Algorithms

Rapid progress has also been made in the area of digital signal processing hardware which might be used to implement the algorithms described in the preceding section. The specific technology which would be utilized depends heavily on

> (1) the number of aperture samples,
>
> (2) the specific algorithm, and
>
> (3) the requisite data rate.

Inasmuch as no specific implementation has been addressed, it is not possible to present a detailed analysis. However, examination of some representative algorithms indicates the type of hardware which may be required.

The number of multiplies and adds (MAD) used by an algorithm is a commonly used index of the degree of numerical difficulty for real-time applica-

6-36

tions. In Table 6-4 we show the number of MAD required to implement the MEM spectral algorithm for a 20-element linear array.

If all of the computations indicated in Table 6-4 were to be accomplished on a contemporary minicomputer (5µs MAD) (e.g., as was done for the DLS test system at Braunschweig), the spectrum AR computation would take approximately 10 msec. Assuming an aircraft data rate of 5 Hz, this would suggest analyzing signals from at most 20 aircraft per second. Consequently, we conclude that special-purpose digital processing hardware would probably be required to make these algorithms useful.

The covariance formation is seen to be an important computational load for either of these algorithms. One effective approach to performing this particular manipulation is the use of hardwired multiplier chips to form the products. For example, there currently are available 16-bit x 16-bit multipliers with accumulate capability for $350 which accomplishes a multiplication in 155 nsec. A bank of 10 such multipliers could form the initial covariance calculation in approximately 15 µsec. The spatial covariance smoothing (approximately 400 complex adds) could be accomplished in the same 15 µsec by pipelined hard-wired adders.

The normal equation solution would be more costly to achieve in a hard-wired implementation this time. If it were to be accomplished on a contemporary minicomputer (5µs MAD), the processing time would be on the order of 250 µsec.

The spectrum formation requires many computations due to the fine search grid necessitated by the high-resolution estimators. There are currently available array processors which could accomplish such a transform in less than 1 msec. For example, there are fixed-point processors (16 bits, 10 butterfly boards) which can accomplish a 1024-point transform in 488 µsec at a cost of approximately $40K.

Root finding should be more efficient than spectrum analysis in this case because:

(1) one knows exactly how many roots will occur whereas the number of spectrum peaks is not known a priori (the

TABLE 6-4

COMPUTATIONS REQUIRED FOR MAXIMUM AUTO REGRESSIVE ALGORITHMS WITH
20-ELEMENT ARRAY AND 6TH ORDER AR MODEL

	Complex Multiplies	Complex Adds
Covariance Formation[1]	210	392
Normal Equation Solution[2]	57	57
Spectrum Formation[3]	1052	2084
Root Finding[4]	216	216

Notes

1. Assumes no time averaging of received data.

2. Using LDU decomposition of Morf et al [49].

3. FFT analysis with pruning and doubling to give frequency resolution to 0.02 beamwidth (1024 points).

4. Estimated to be $\sim N_p^{1.74}$ where N_p = number of zeros.

number of spectrum peaks is upper bounded by the number of roots), and

(2) the results of other estimators (e.g., beamsum) can be used to furnish initial estimates for some of the roots.

Analytical expressions for the required number of computations for finding the roots of a complex polynomial are not available. Figure 6-12 shows the results of our timing experiments on the Lincoln Laboratory Amdahl 470 computer timeshare system for a sample case involving two coherent signals at 0.5 BW separation angle, 45° rf phase, and 40 dB array SNR impinging on a 20-element array.[*] The ARM root computations are seen to require approximately 10 msec for a 7th-order AR model. It is difficult to assign a single MAD execution time for this computer in the time share mode since

(1) the fetch time from memory will depend on whether the location is within the current in-core "page" as opposed to a disk, and

(2) register storage of variables can substantially reduce the fetch time required to well below that required for core memory.

A "same page" memory access from core would yield a 7-µs MAD for the double-precision complex variables used in our current root finding routine, which corresponds reasonably well to the 5-µs MAD we used in the above computation time estimates.

On the other hand, it is clear from Fig. 6-12 that cer. 'r of the routines (especially, the normal equation solution) took much lot.er (relatively) than would have been predicted from Table 6-4. This type of difference illustrates the difficulties that can arise when a non-optimized higher order programming language (e.g., Fortran) is used for computation-intensive algorithms. More detailed studies using optimized assembly language code are required, but could not be accomplished in the current study.

---

[*] It should be stressed that none of the algorithms were particularly optimized for a computational speed (e.g., Fortran code was used).

Fig. 6-12a. Timing estimates for maximum entropy algorithms.

two signals

all computations double precision complex

ensemble (correlated)
covariance

——— ZCPXLY
········· CRLS
—·—·— SPECTRUM
—–—–— MCOV

normal eqn.
solution

root finding

covariance estimation
and smoothing

400 point spectrum
computation

TIME IN SEC

MODEL ORDER (LOG)

10^0 10^{-1} 10^{-2} 10^{-3} 10^{-4}

10^0 10^1 10^2
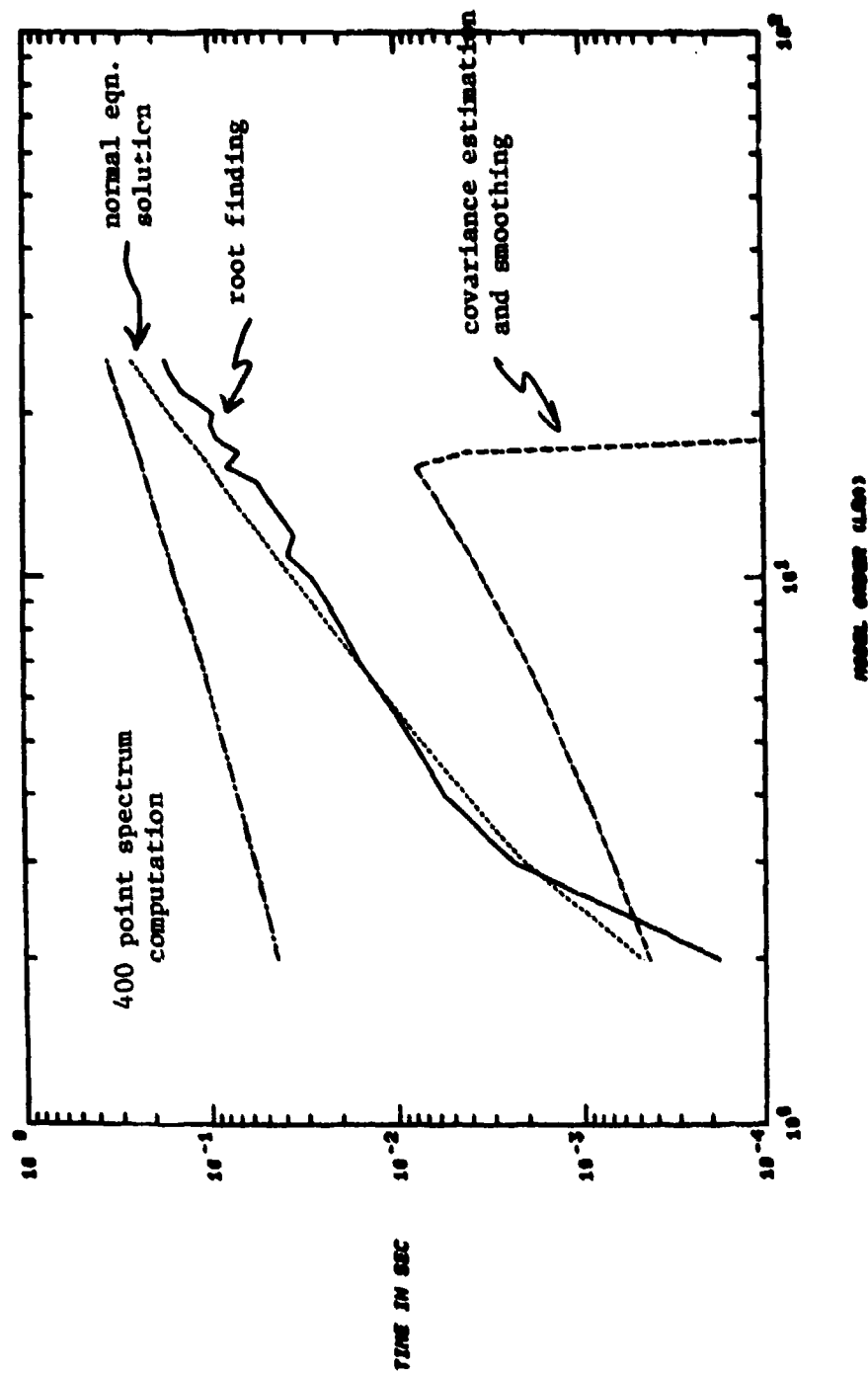
Fig. 6-12b. Timing estimates for maximum entropy analysis.

6-41
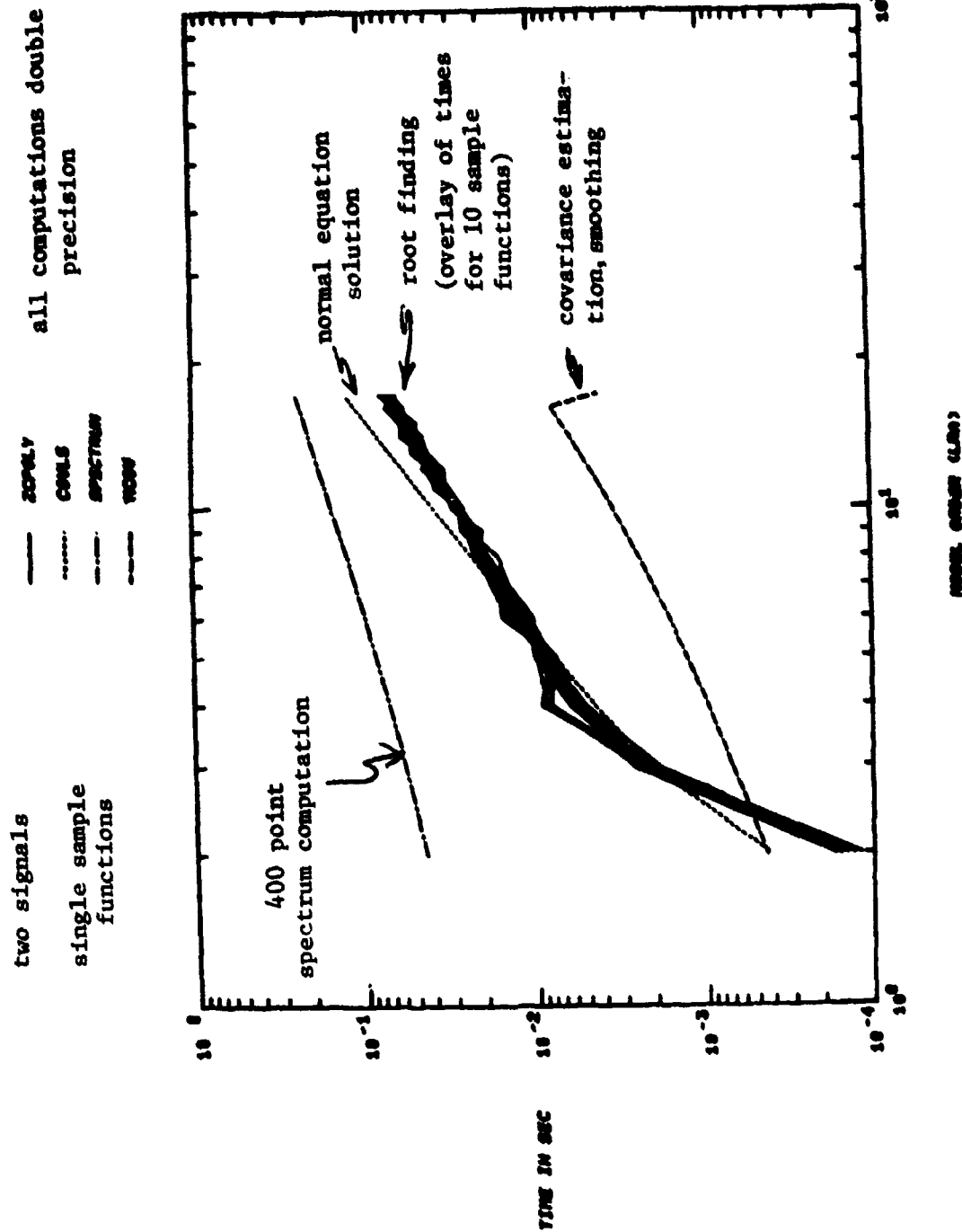
Fig. 6-12c. Timing estimates for maximum entropy analysis.
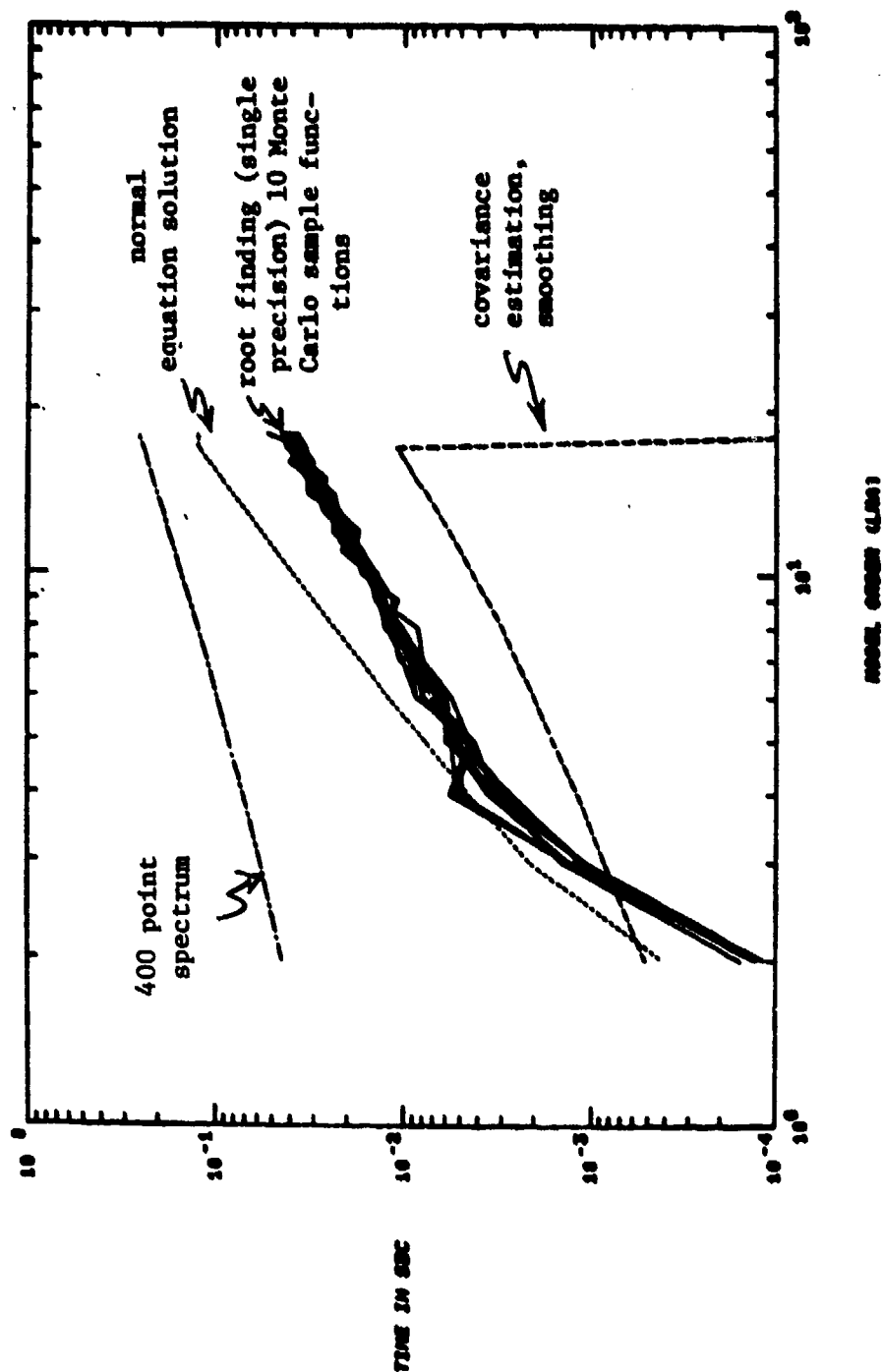
Fig. 6-12d. Timing estimates for maximum entropy analysis.

Computations of the required complex operations were also accomplished for the MLM technique using the same degree of covariance smoothing. For MLM, the normal equation solution is replaced by a matrix inversion which requires approximately 7 times as many complex MAD as were required for the AR estimator. Also, with the MLM one would have to perform a spectrum computation as opposed to root finding to determine the plane wave angles.

VII. APPLICATION OF APERTURE SAMPLING ALGORITHMS TO MLS ANGLE GUIDANCE SIGNAL PROCESSING

In this section, we consider how the various aperture sampling algorithms discussed in the preceding chapters might be applied to MLS angle process- ing. It should be emphasized at the outset that the results here are less detailed than some of the preceding sections since our prime objective was to identify the principal problems in adapting the high-resolution techniques to the MLS angle systems.

Figure 7-1 shows how the MLS angle guidance is furnished by measuring the true difference between passages of a scanning beam past the receiver (the time reference scanning beam (TRSB) technique). Also illustrated in Fig. 7-1 is how the current TRSB receiver determines the beam centroids by thresholding on the scanning-beam envelope.

Figure 4-1 showed how this type of processing combined with the common form of TRSB scanning beam generation via a phased array is in fact analogous to Fourier analysis in a sampled-aperture system. Unfortunately, the observed quantity in the current TRSB receivers is typically the log magnitude of the Fourier transform, i.e., the phase information has been lost. If we were able to recover the phase information, then by inverse Fourier transform of the (now) complex envelope one could obtain the sampled aperture data which was the starting point for the algorithms of Chapters III-V.

Two approaches to phase information recovery are considered in this chapter. The first involves complex demodulation of the scanning-beam en- velope using the DPSK signal which is radiated prior to the beam scan peri- od. The second approach is to recover the sampled-aperture data from the envelope magnitude data alone. Section C considers the potential improvement which might be afforded by the techniques discussed earlier for elevation angle estimation vis-a-vis the performance offered by the single edge proces- sor (SEP).

A. Complex Demodulation of Scanning-Beam Signals

The TRSB MLS signal format (Figs. 7-2 and 7-3) includes radiation of a DPSK preamble signal of 1.6 msec duration prior to the beam scan period. The
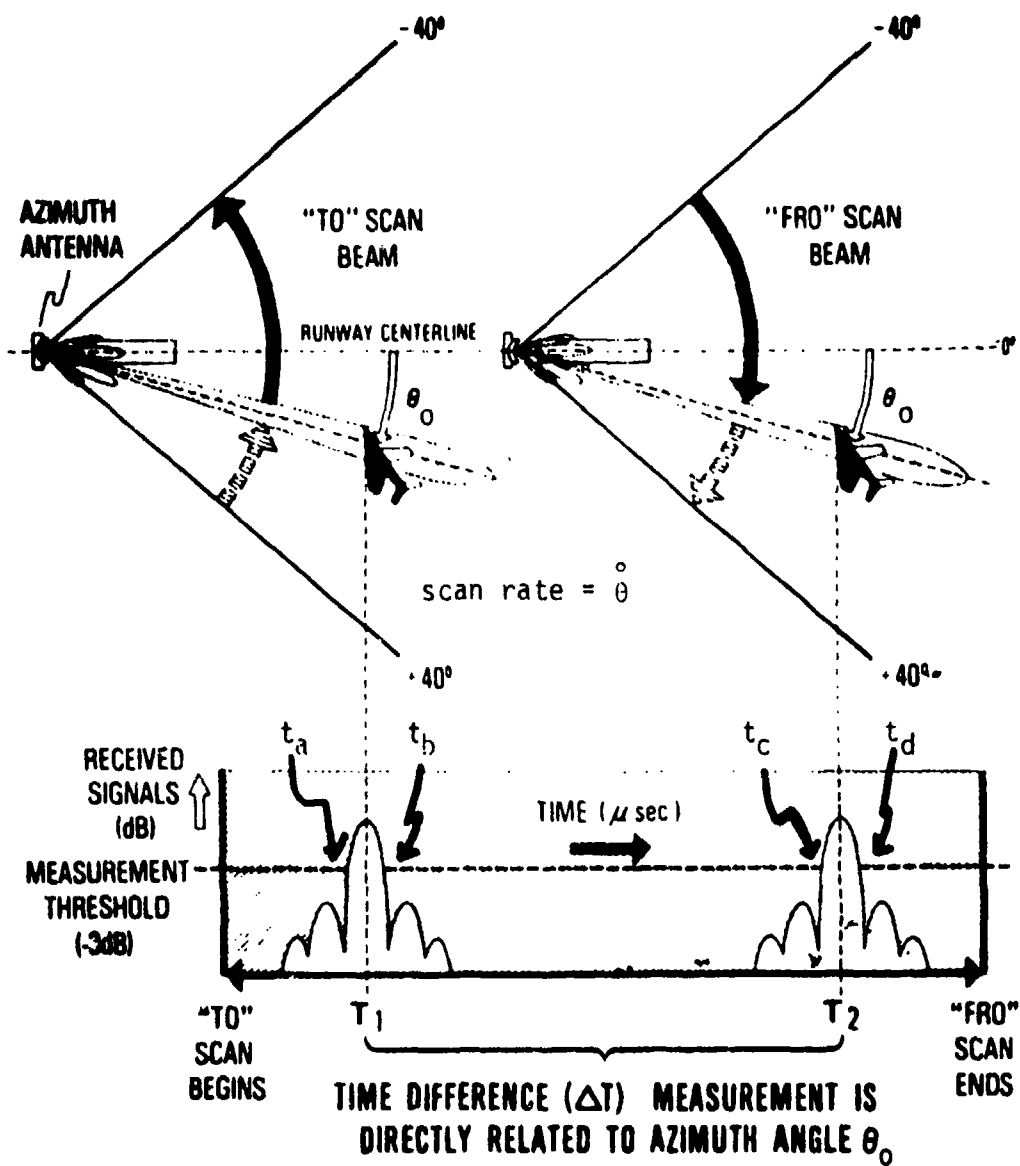
Fig. 7-1. TRSB angle measurement.
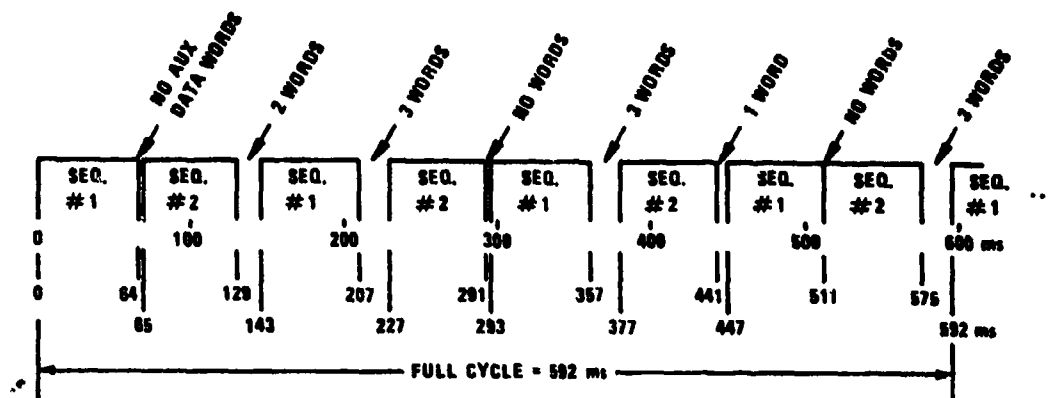
(a) Full Cycle of Functions
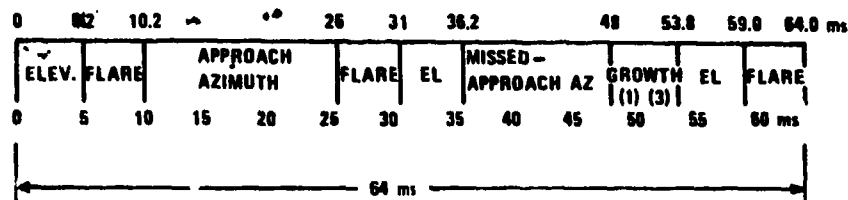
(b) Subsequence #1
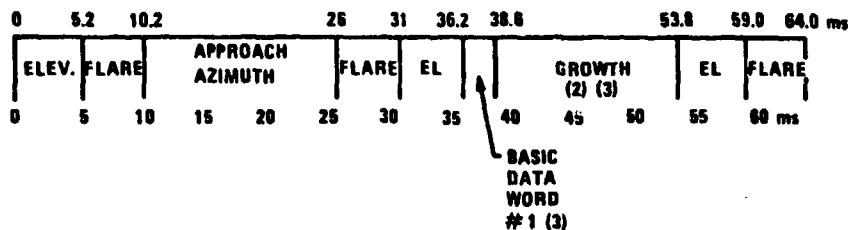
(c) Subsequence #2

NOTES: (1) AUXILIARY DATA (1 WORD) OR MISSED APPROACH ELEVATION
      (2) 360° AZIMUTH OR AUXILIARY DATA (2 WORDS)
      (3) BASIC DATA WORD #2 TRANSMITTED EITHER IN GROWTH
          OR BETWEEN SEQUENCES

Fig. 7-2. TRSB signal format (TDM).

* Basic Narrow scan limits illustrated; small community limits are
  ±12° while expanded limits are ±62°

Fly Rt, Fly Left pulses present only with small community system
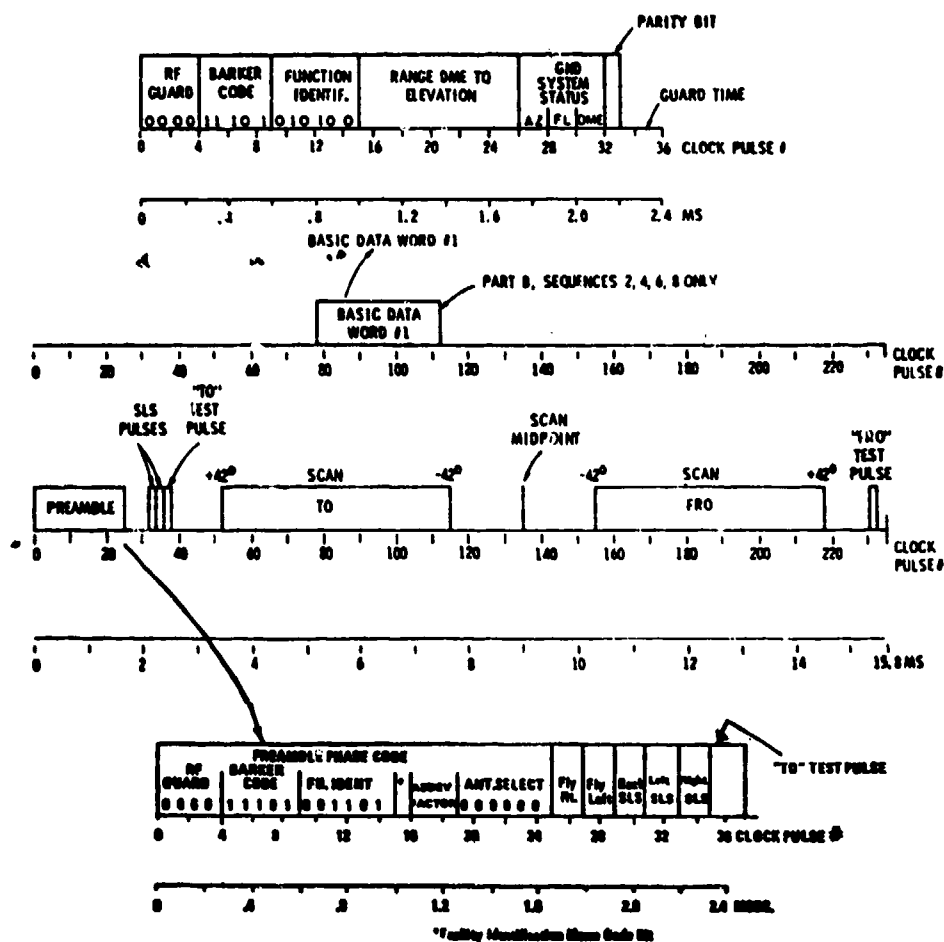
Fig. 7-3a. Azimuth signal format.
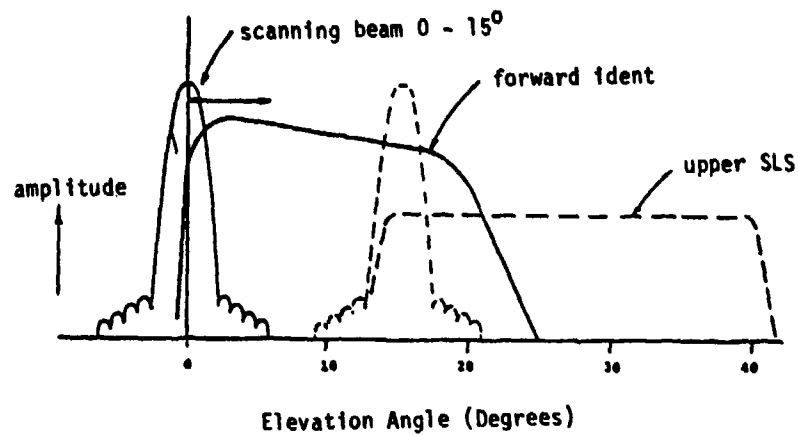
ELEVATION COVERAGE

Signal Format
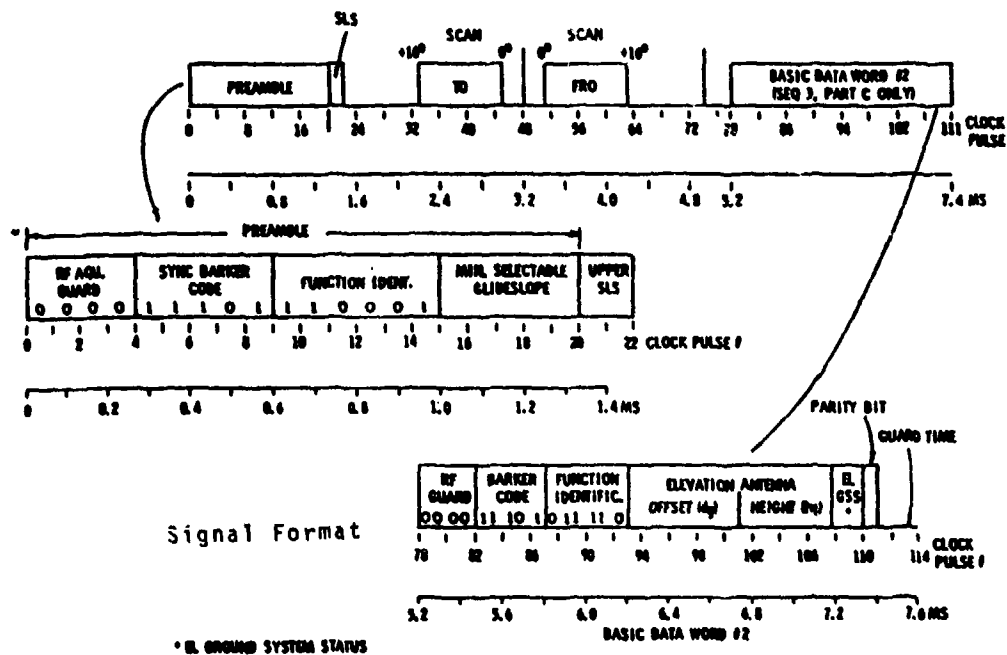
Fig. 7-3b. Elevation function coverage and format.

DPSK antenna uses the same transmitter (typically a TWT) and local oscillator that is used to furnish rf power to the phased array. Thus, the possibility then arises to lock a reference oscillator in the aircraft receiver to the DPSK carrier and then use the reference oscillator signal to coherently de-modulate the scanning-beam signal.

The first question that arises with such a scheme is the accuracy to which the DPSK carrier frequency can be measured and what will be the effects of a frequency offset between the actual carrier frequency and the aircraft reference frequency. In the simplest implementation, the reference oscillator locks onto the DPSK carrier anew on each scan. In this case, the carrier frequency rms error is lower bounded by

$$\delta_f = \frac{\sqrt{3}}{\pi \ T} \ \ \frac{1}{\sqrt{SNR}} \qquad\qquad (7\text{-}1)$$

$$= 550 \ \text{Hz}$$

for T = 1.6 msec and a SNR = 100 (20 dB). If the DPSK carrier frequency can be measured over N function preambles, δ_f would be reduced by a factor of $\sqrt{N}$. * In either case, (based on the Lincoln multipath measurement exper-ience [5]) the uncertainty in δ_f due to the DPSK short duration should be several times larger than the local oscillator drifts between the DPSK and scanning-beam transmissions.

Another potentially major source of carrier frequency offsets is the "reference scalloping" effect due to specular multipath. This error, which arises when the receiver-motion-induced Doppler frequency for the multipath differs significantly from that for the direct path, was examined in detail in the context of the Doppler MLS (DMLS) [28, 29]. The error for the case consi-dered here corresponds fairly closely to the DMLS situation with a single antenna scan per data frame. The carrier frequency error depends on the

_____

*This averaging over successive receptions of a given function (e.g. azimuth) which are interspersed between receptions of other functions (e.g., elevation) might be difficult for an analog implementation.

relative multipath level (ρ) and the difference, f_s, in the relative Doppler frequencies, and could approach 500 Hz for a M/D level of -6 dB. For low frequency differences, the error is upper bounded by the product ρf_s. Thus, terrain reflections and/or diffraction multipath would not cause sizable frequency errors since f_s is typically small for these cases.

Next, we want to consider the effect of a frequency error, ω, on the estimated aperture sample data. Let the received signal along the aperture be $s(x)$. Then the received complex TRSB envelope in the absence of noise is

$$V(t) = \sum_i s(x_i) \, e^{jx_i \sin\dot\theta t} \approx \int s(x) \, e^{jx \sin\dot\theta t} d\,x \,. \qquad (7\text{-}2)$$

The estimated aperture distribution is then

$$\hat s(x) = \int \dot\theta V(t) \, e^{-jx \sin\dot\theta t} \cos\dot\theta t \, dt \qquad (7\text{-}3)$$

where various normalizing constants have been ignored. When there is a frequency error δw

$$\hat s(x) = \dot\theta \int V(t) \, e^{j\delta wt} e^{-jx\sin\dot\theta t} \cos \dot\theta t \, dt \qquad (7\text{-}4)$$

If we assume that the principal contribution of $V(t)$ is near boresight ($\dot\theta t \approx 0$), then

$$\hat s(x) \sim \dot\theta \int V(t) \, e^{-j(\theta x - \delta w)t} dt \qquad (7\text{-}5)$$
$$\approx s(x - x_o)$$

where $x_o = \delta w / \dot\theta$ is the equivalent aperture displacement in wavelengths. Using the numerical values

$$\dot\theta = 20000 \text{ deg/sec} = 349 \text{ rad/sec}$$

$$\delta w = 2\pi \cdot \left(550\right) = 3455 \text{ rad/sec}$$

we obtain

$$x_o = 0.1\lambda .$$

This (small) shift in array origin will change the relative phases of the various signals by an amount $2\pi x_o \sin\theta_1$, but otherwise would not change the ability to analyze the angles of arrival. Equation (7-4) can be rewritten as an integral over $\eta = \sin\dot\theta t$:

$$\hat{s}(x) = \dot\theta \int V(\eta) \exp(j \frac{\delta w}{\dot\theta} \sin^{-1}\eta - \eta x) \, d\eta \qquad (7-6)$$

We have not been able to reduce this integral to a closed form by the usual manipulations. Numerical integration would be relatively straightforward, but was not accomplished during the current phase of the study. It might be noted that had ICAO chosen a MLS scan format which was linear in η so that the phase shifters could step uniformly in time, then equation (7-5) would be rigorously true.

Assuming that one has successfully recovered the TRSB envelope phase by one means or the other, we now want to consider some realization issues. The shift of time scale to sample $V(t)$ uniformly in η as opposed to t is relatively straightforward. The number of envelope samples required is not quite clear due to the following:

    (i) if there were no front end noise, then only N equally spaced (in η) envelope samples are required to determine the equivalent aperture wavefront at the N elements of the ground array.

    (ii) if the bandwidth of the IF filter is W, then the noise samples spaced approximately 1/2W seconds apart are uncorrelated (and probably independent). The N-point transform of such samples would yield N uncorrelated samples which is analogous to independent noise at N front ends of an aperture sampling array.

Thus, the equal spacing in η (= $\sin\dot\theta t$) suggested by (i) seems incompatible with the equal spacing in t suggested by noise considerations. If one needs only to sample near boresight of the array (e.g., direct signal near boresight

7-6

and a wide IF filter bandwidth), the two sampling grids are essentially identical. The general case would (as in the case of a frequency offset) require numerical assessment.

Another practical difficulty is the apparent necessity to know the number of antenna elements, spacing, and aperture distribution. These quantities are not a part of the MLS data transmitted to the receiver although the nominal ground-antenna beamwidth and proportional coverage limits are known. The coverage limits and need to avoid grating lobes in the forward sector set an upper limit on possible element spacings. Arrays are normally designed with a spacing near this upper limit so as to minimize costs. Similarly, given the beamwidth, one can make a fairly reasonable guess as to the total aperture length (and hence, number of elements). We suspect, but have not yet proved, that errors of 10% in number of elements and spacings may be acceptable.

The issue of aperture distribution is more difficult. Most high-efficiency, low-sidelobe patterns have roughly similar normalized aperture distributions [73]. Thus, a nominal correction may suffice to remove some of these effects. However, there are some knotty theoretical issues associated with the need to avoid increasing the noise contribution significantly for the wavefront samples which were heavily altered by the aperture distribution.* The degree to which noise power can be traded off with wavefront amplitude errors due to the (not fully compensated) aperture weighting has not been examined quantitatively.

Provided that the above difficulties can be surmounted, the actual Fourier inversion of the complex envelope would be relatively straightforward to realize with the microprocessors which will be available in the near future. This is based on the observation that 1) the number of scans/sec for the TRSB functions is approximately 1/2 the number that arose with the DMLS

---

*Note: the noise here is added on after the aperture distribution has been applied, whereas with a ground-derived system one would alter both the received wavefront and the front-end noise in weighting the received data to achieve the desired sidelobes.

receivers [28] and 2) the DMLS receivers did not utilize FFT techniques to accomplish their Fourier transforms.

B.   Reconstruction of the Aperture Data from the Envelope Magnitude

In general, a signal cannot be uniquely specified by only the phase or magnitude of its Fourier transform. However, one condition under which the magnitude and phase are related is the minimum-phase condition and under this condition a signal can be uniquely recovered from the magnitude of its Fourier transform, or, to within a scale factor, the phase of its Fourier transform.

Conditions under which a signal is minimum phase and procedures for accomplishing the inversion have been considered by Quatieri and Oppenheim [74]. In particular, they consider the case where the z-transform of a sequence $\{r_i\}$ to be a rational function which we express in the form

$$R(z) = Az^{n_o} \frac{\prod_{k=1}^{M_i} (1-a_k z^{-1}) \prod_{k=1}^{M_o} (1-b_k z)}{\prod_{k=1}^{P_i} (1-c_k z^{-1}) \prod_{k=1}^{P_o} (1-d_k z)} \qquad (7-7)$$

where $|a_k|$, $|b_k|$, $|c_k|$, and $|d_k|$ are less than or equal to unity, z^{n_o} is a linear phase factor, and A is a scale factor. When, in addition, $\{r_i\}$ is stable, i.e., $\sum_i |r_i| < \infty$, then $|c_k|$ and $|d_k|$ are strictly less than one.

R(z) is minimum phase if it and its reciprocal $R^{-1}(z)$ are both analytic for $|z| \geq 1$. Quatieri and Oppenheim show that this implies that n_o, $\{b_k\}$, and $\{d_k\}$ are zero in Eq. (7-7) and the zero and poles of R(z) lie within the unit circle. This in turn implies that the received signal can be represented by an ARMA model.

Another condition which can be shown to be equivalent to the above definition of a minimum-phase sequence is that the log-magnitude and unwrapped phase of R(z) are related through the Hilbert transform [76]. The Hilbert transform relation guarantees that a minimum-phase sequence can be uniquely

7-10

specified from the Fourier transform phase and, to within a scale factor, from the Fourier transform magnitude.

For a minimum-phase signal, the log-magnitude and phase of the Fourier transform are related through the Hilbert transform and the direct implementation of the Hilbert transform using the DFT has been extensively investigated [75,77]. One disadvantage of this implementation is that in computing the phase from the log-magnitude, estimates of the unwrapped phase are required and are often difficult to compute. A second drawback is that aliasing occurs in the inverse discrete Fourier transform of samples of the log-magnitude and unwrapped phase due to a finite DFT length, limiting the accuracy of the computed samples of the unknown component.

Quatieri and Oppenheim [74] suggest an iterative algorithm for reconstructing the minimum-phase sequence $\{r_i\}$ from the magnitude of its Fourier transform by imposing, in an iterative fashion, causality and the initial value r_0 in the space domain, and the known magnitude in the z domain. When the algorithm results in a sequence which satisfies certain minimum-phase conditions and has the given magnitude, it must equal $\{\tilde{r}_i\}$. The iterative procedure involves repeated Fourier transforms and inverse transforms.

Although successful convergence of the iterative procedure was accomplished by Quatieri and Oppenheim, the convergence was sometimes slow (e.g., several hundred iterations) and the sensitivity to inexactness in the log-magnitude, quantization noise is not understood.

Although the received waveform $\{s_i\}$ probably can be represented by a model such as Eq. (7-7), it is not clear whether additive noise would yield a minimum-phase $\{r_i\}$. This problem together with the significant computational cost and issues for the reconstruction techniques has lead us to conclude that these techniques do not appear to be practically usable in an MLS receiver in the near future.

C.   Performance Improvement With High-Resolution Sampling Algorithms for
     MLS Elevation Angle Estimation

In the preceding sections, we have indicated that there are some diffi-
culties in recovering sampled-aperture data for the TRSB due to the signal
format and current receiver processing. On the other hand, one can consider
whether there are situations where improved angle measurement performance in
the presence of multipath would be of use. The most demanding requirements
for MLS are in the final approach and landing phase. In this region, the
ground-antenna siting and beamwidths are typically chosen so that all of the
multipath signals are out of beam (i.e., resolvable by the Rayleigh criteria)
except for:

    (i) diffraction multipath due to taxiing or overflying
         aircraft and fixed objects (e.g., monitor poles),

    (ii) terrain reflection multipath for the flare system, and

    (iii)building wall reflections of the elevation signal.

Each of the above has been successfully coped with in the field tests and
scenario simulations to date [28, 29, 80]. However, it is probably fair to
state that the flare terrain multipath problem has 1) received the fewest
field tests [86] (due to lack of a suitable flare antenna and signal processor
during much of the test program) and 2) represents the most difficult chal-
lenge in terms of long-duration, nigh-level multipath with a small separation
angle [85]. Similar elevation multipath problems may arise when using small-
aperture elevation arrays (e.g., as for military tactical or civilian "emer-
gency" use) to provide guidance at low elevation angles (e.g., < 3°) over
rising and/or rolling terrain.

Thus, it is useful to compare the performance of the single edge proces-
sor (SEP) used as a TRSB low elevation angle/flare processor [79, 85] with the
performance of the extended Pisarenko/AR estimator (ME) discussed in Chapter
IV. The SEP can be viewed as attempting to improve the estimate of the direct
signal angle by analyzing the Fourier transform (i.e., beamsum) peak on the

side "away from" the multipath[*] (see Fig. 7-4). Also considered were several other radar elevation tracking algorithms which attempt to take advantage of the elevation multipath angular location:

(a) conventional null seeking monopulse (CM) systems can be viewed as determining the peak of the beamsum (BS) power spectrum $|\Sigma(\theta)|^2$ since the null of the ratio $Re[\Delta(\theta)/\Sigma(\theta)]$ with difference pattern $\Delta(\theta) = d\Sigma(\theta)/d\theta$ occurs at the peak of the beamsum pattern $\Sigma(\theta)$. The principal difference between monopulse and the beamsum error response to inbeam multipath occurs at high M/D levels with 1) large separation angles with a relative phase near $0°$ and 2) at all separation angles with a relative phase near $\pm 90°$ [28].

(b) the off-boresight monopulse (OBM) technique [78] which utilizes the fact that when only a direct signal is present at angle θ_d and the antenna boresight angle θ is within 1 beamwidth of θ_d,

$$\varepsilon(\theta) \triangleq Re[\Delta(\theta)/\Sigma(\theta)] = (N\delta)\ (\theta-\theta_d) \qquad (7-8)$$

so that one can estimate θ_d without pointing the array at θ_d. With an "off-boresight" elevation tracker θ is constrained to be $> 0.7/N\delta$ and Eq. (7-8) is used to estimate θ_d if the last estimate of θ_d is less than $0.7/N\delta$. This keeps the main lobes of $\Sigma(\theta)$ and $\Delta(\theta)$ pointed above the terrain and thus significantly reduces the errors due to multipath signals at elevation angles below θ_d [78].

(c) At low elevation angles, it may not be possible to prevent the mainlobes of $\Delta(\theta)$ from illuminating the ground. W. White [18] has suggested use of a double null monopulse (DNM) in which the difference pattern has a null both at θ and at the expected angle of the multipath $(=-\theta)$. The error metric $\varepsilon(\theta)$ for White's monopulse estimator has a null at $\theta=\theta_d$ when only the direct signal is present and when the multipath is present at $-\theta_d$.

Figure 7-5 illustrates the various angle determination techniques in the context of a ground-derived system.

---

[*]This "one sided" processing (analogous to leading-edge range estimation in radar and DME systems [44]) is plausible here because the terrain multipath is known to be at a negative elevation angle with respect to the direct signal angle.

(a) DWELL GATE PROCESSOR

t_1 AND t_2 FOUND BY INTERPOLATION
BETWEEN SAMPLING POINTS
(dashed lines)

DWELL GATE WIDTH $= t_2 - t_1$

BEAM CENTER TIME $= \dfrac{t_1 + t_2}{2}$

3 dB

PEAK
DIGITZED
VALUE

t_n t_1 t_{n+1} t_2 t_m t_{m+1}
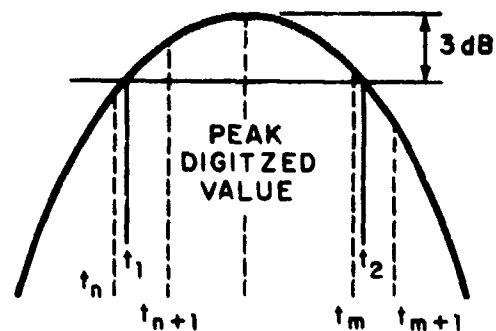
(b) SINGLE EDGE PROCESSOR

TIME OF 9 dB IN 16 μs SLOPE
FOUND BY INTERPOLATION
BETWEEN TIMES OF GREATER
AND LESS SLOPE (Equivalent to analog
delay and compare thresholding)

9 dB
9 dB
16 μs
t_n t_{n+3}
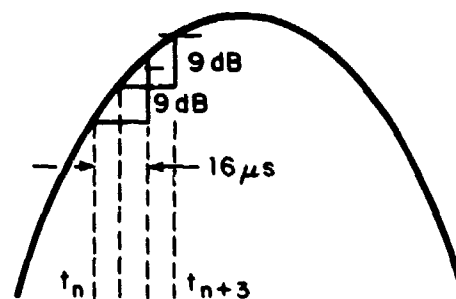
(c) SPLIT GATE PROCESSOR

DIFFERENCE OF SUMS OF FOUR
AMPLITUDES ON EITHER SIDE OF
PEAK $(\Sigma_1^+ - \Sigma_1^-)$ IS INTERPOLATED
WITH SHIFTED DIFFERENCE $(\Sigma_2^+ - \Sigma_2^-)$
TO DETERMINE BEAM CENTROID

PEAK VALUE

Σ_1^- Σ_1^+

Σ_2^- Σ_2^+
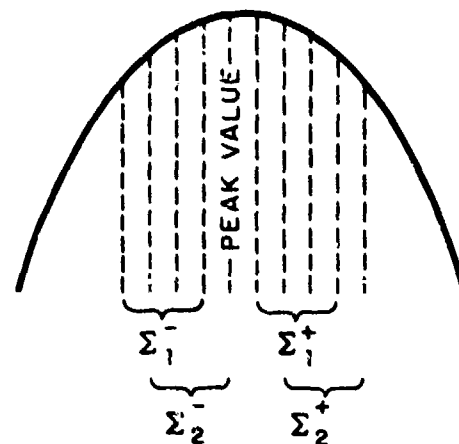
Fig. 7-4. Angle processing techniques studied by Calspan (79).

TARGET ELEVATION ANGLE ESTIMATORS

CONVENTIONAL MONOPULSE

OFF-BORESIGHT MONOPULSE

DOUBLE-NULL PROCESSING

SINGLE-EDGE PROCESSING

SAMPLED APERTURE PROCESSING

A/D CONVERTER

COMPUTER
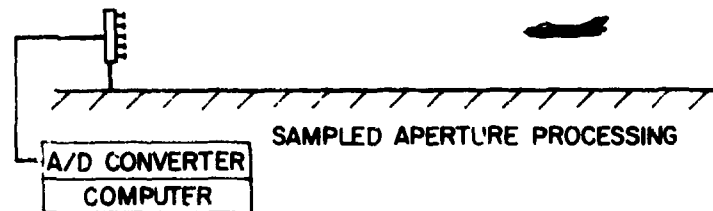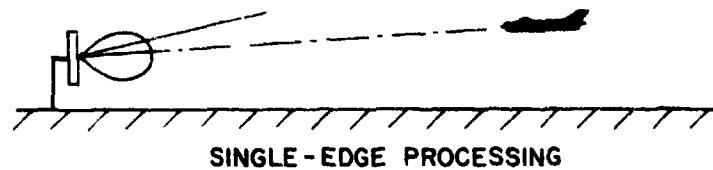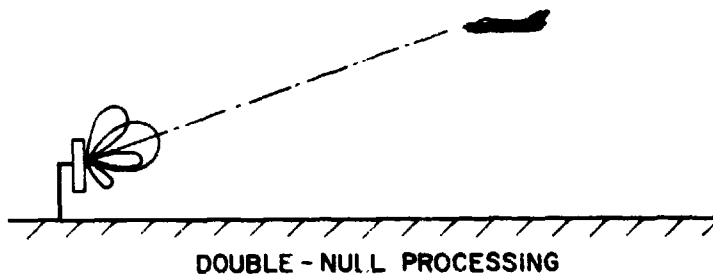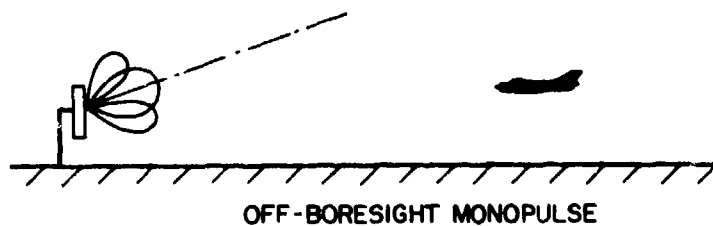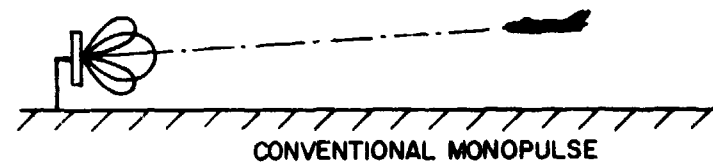
Fig. 7-5. Target elevation angle estimators.

7-15

1. Simulation Results

Figure 7-6 shows the simulation results[*] for the various angle estimation techniques for a single multipath signal of relative amplitude $\rho = 0.9$ at angle θ_m when the direct signal is at $\theta_d = 0.5$ BW, with 30-dB element SNR. Both $0°$ and $180°$ relative phase conditions are shown since these two bound the errors at other relative phases. The ME technique appears to give the best performance over the range of angles with the SEP and DNM techniques a close second.

2. Experimental Results

L- and C-band terrain multipath field measurement data obtained with the experimental setup of Fig. 4-44 were used to compare the various estimators response to inbeam reflection multipath. The C-band results were obtained with a 10-element 14.3λ array, while the L-band results were obtained with a 5-element 6.5λ array.[**] The angular error shown in the following figures is the difference between the estimated angle and the theodolite tracking angle at that particular moment. Therefore, it is understood that the angular error also includes the possible theodolite tracking error. The theodolite tracking error is expected to be on the order of $0.1°$ for the measurements described here.

Figure 7-7 shows the errors for the various techniques as a function of elevation angle for a test at L.G. Hanscom Airport, Massachusetts. The terrain here is nominally flat with grass cover. We see that in general all techniques except the SEP have similar performance for $\theta > 1$ beamwidth (BW), where θ_d is the target helicopter elevation angle. The maximum angular error

---

[*]All techniques were simulated with a 40-45 dB array SNR.

[**]The small number of L-band array elements and limited number of time observations necessitated using a two-pole model for the ME technique, whereas the observed angular spectra using 9-element arrays would in several cases require at least a four-pole model. Consequently, the ME L-band errors are greater than would have been the case with a greater number of elements within the overall 6.5λ aperture.
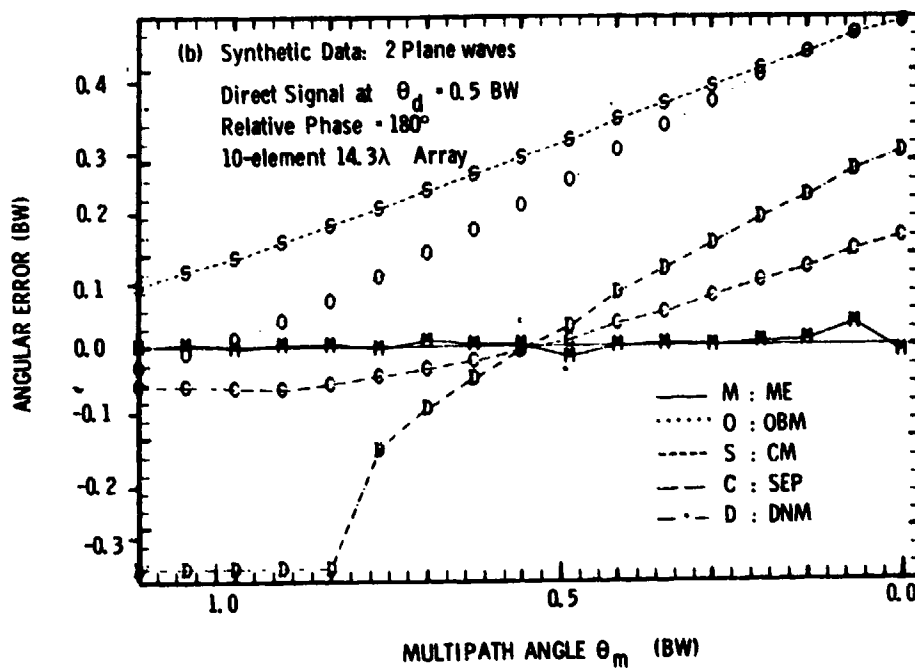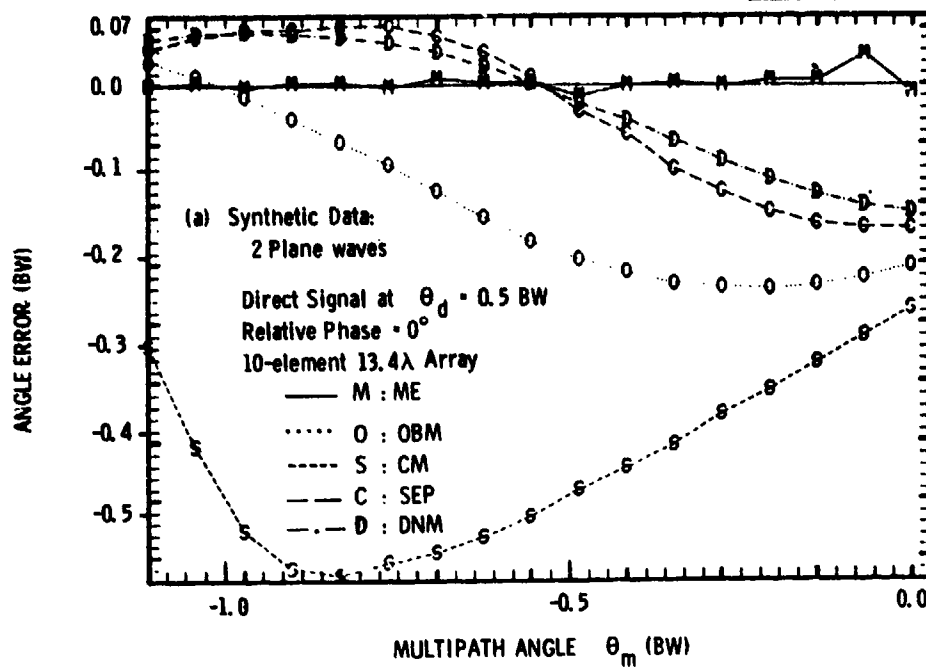
Fig. 7-6. Synthetic data case.
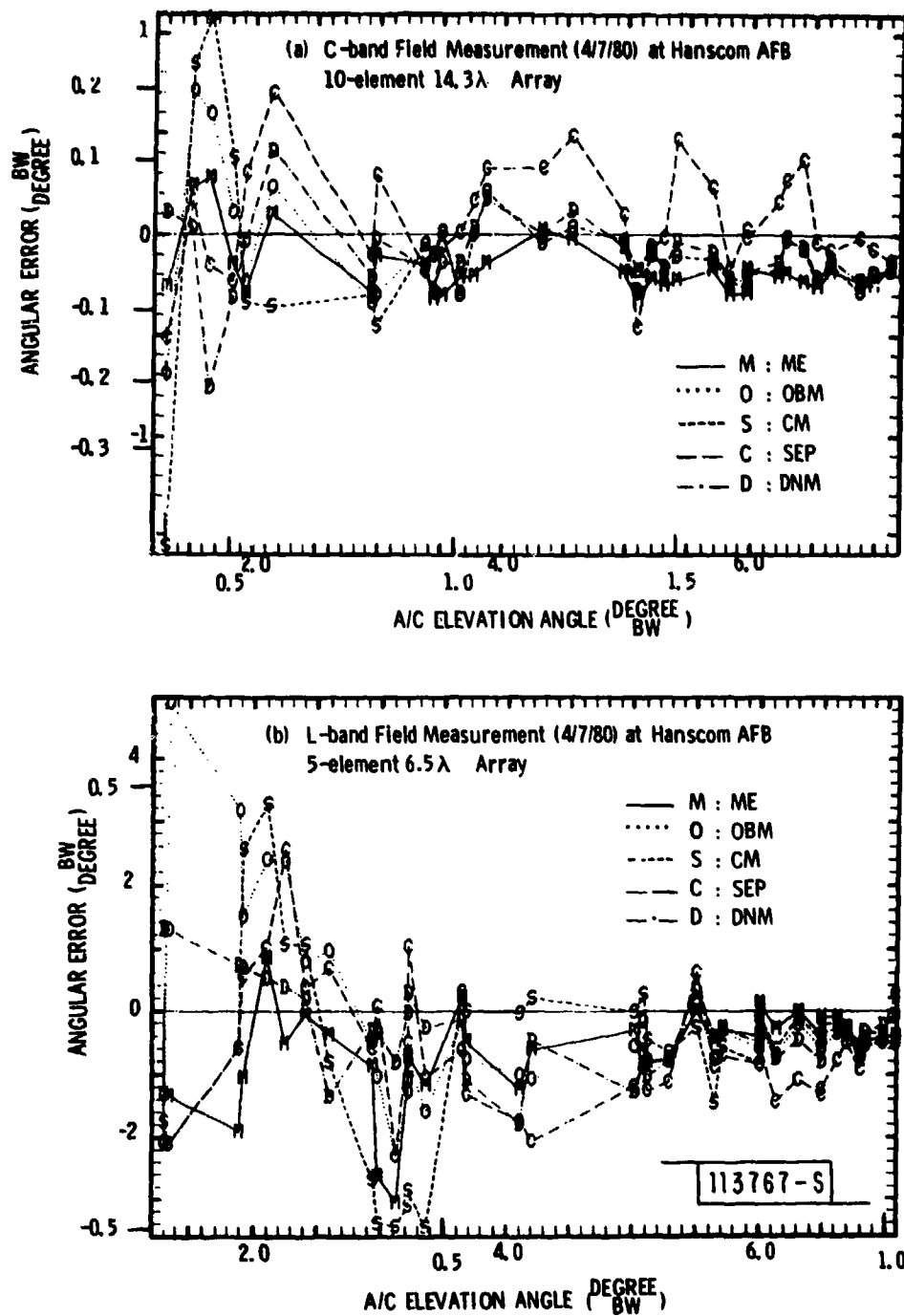
Fig. 7-7.   Hanscom AFB measurement: near-flat terrain.

7-18

($\Delta\theta$ max) is about 0.07 BW, except for the SEP which shows larger error of 0.12 BW. For 1 BW > θ_d >0.5 BW, it appears that the ME, DNM, and OBM techniques yield smaller error ($\Delta\theta$ max ~ 0.08 BW), with the ME technique having the best performance around θ_d ~ 0.5 BW. For θ_d < 0.5 BW, the ME technique appears to give the best performance with the angular error similar to that observed in the higher elevation angles. The larger angular error in the L-band results (Fig. 7-7b), especially with the ME technique, probably is due to the insufficient sensor samples (only 5 signal samples available for the L-band versus 10 for the C-band).

Figure 7-8 shows the errors for the various techniques as a function of elevation angle for a test at the Ft. Devens, Massachusetts, golf course. This terrain is rolling with closely cropped grass, as to yield specular reflections at several elevation angles. In general, the angular errors are larger for the rolling terrain here than those previously observed for the near-flat terrain. Also, the small number of sensor samples has a more pronounced effect on the L-band angle estimation accuracy here, especially with the ME technique, than previously observed for the near-flat terrain. This is thought reasonable, since the multipath environment was found to be more complicated for this rolling terrain than for the near-flat terrain. For the C-band result (Fig. 7-8a) which was obtained with a larger number of sensor samples, we can see that again the ME technique appears to yield the best performance, especially for θ_d < 1 BW. The ME angular errors ($\Delta\theta$ max) are around 0.07 to 0.1 BW, except for one isolated elevation angle (θ_d ~ 4.0°) where all techniques show a large angular error (~ 0.15 BW). Again the SEP gives much larger angular error than the other techniques, as we previously observed. This greater error for SEP is believed to arise from the SEP sensitivity to sidelobe multipath.

These preliminary results from applying five different elevation angle estimation techniques to several identical data sets, both synthetic and field measured, indicated that the ME technique based estimator seemed to yield the best performance if a sufficient number of sensor samples was available. The observed maximum angular errors were around 0.07 to 0.1 beamwidth for the target elevation angles from 2 beamwidths down to about 0.3 beamwidth. The
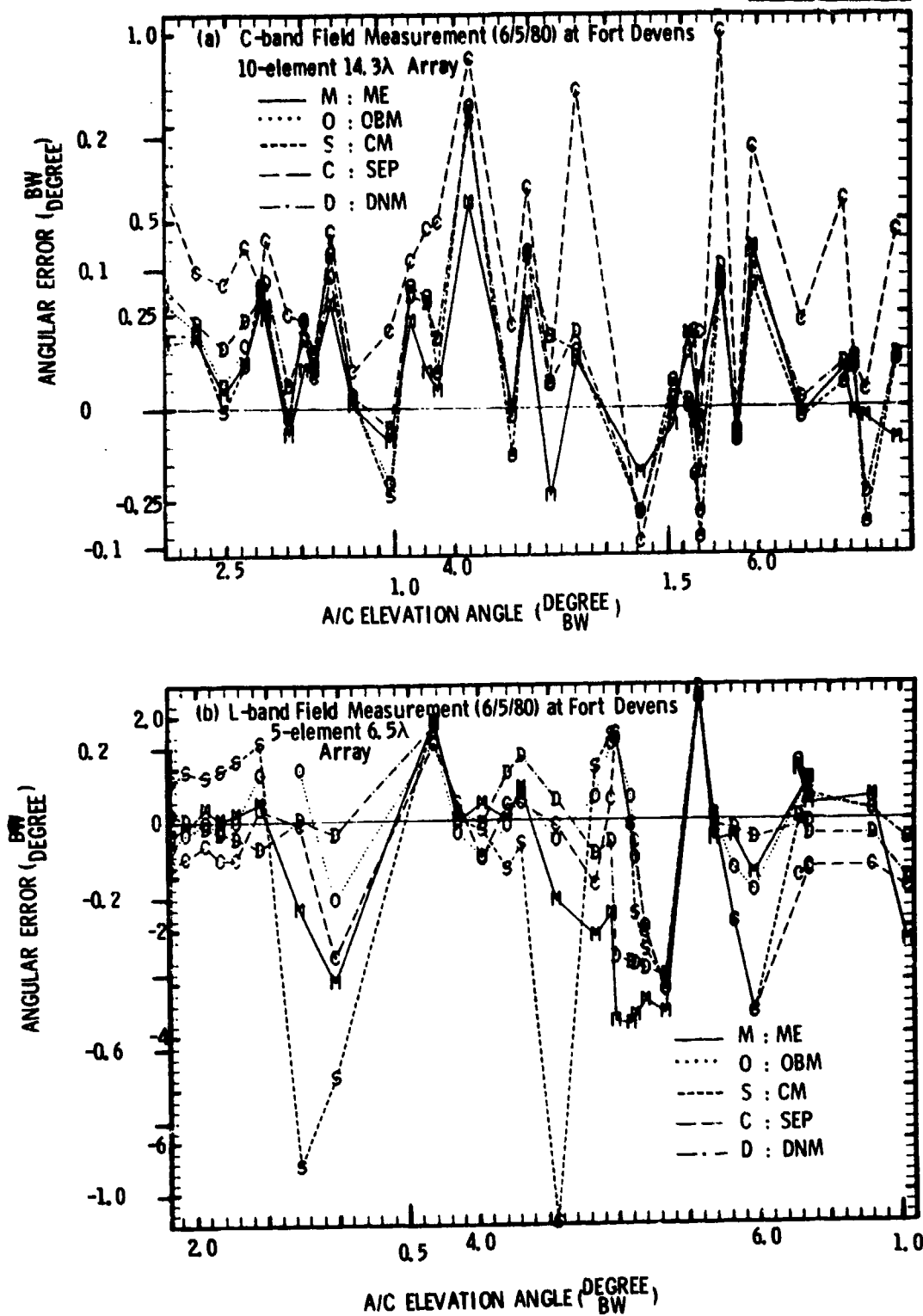
Fig. 7-8. Fort Devens golf course measurement: rolling terrain.

DNM and SEP appeared to work much better than the CM or OBM for the synthetic data cases. However, in the field measurement results, the DNM and OBM seemed to give similar performance and the SEP performed notably poorer than the other elevation angle estimation techniques.

We would caution, however, against drawing any substantive conclusions from these very limited experimental results to the actual MLS situation due to:

(1) the small number of aperture samples available,

(2) the difficulty in carrying out the measurements at elevation angles representative of actual flare situations [85], and

(3) the unrepresentative nature of the golf course terrain as far as flare guidance is concerned.

Rather, we view the experimental data as stressing the possible pitfalls in placing too much credibility of the classical homogeneous flat terrain reflection model for characterizing elevation angle estimation performance with inbeam multipath from irregular terrain.

D. Summary of Results

In this section, we have considered how the advanced array processing techniques discussed in Chapters II - V might be applied to the MLS angle function signal processing. Although the TRSB signal in space is the Fourier transform of the aperture wavefront, we found there were significant issues concerned with

(1) recovering the phase information which is lost in the current TRSB receivers, and

(2) inverse Fourier transform of the received data arising from conflicting receiver noise and scan format factors as well as lack of knowledge as to the element number, spacing, and aperture weighting.

The detailed simulation studies which would be needed to quantify the practical import of these problems could not be accomplished within the scope of the current study. Similarly, although a minimum-phase estimate of the (weighted)

aperture wavefront can be obtained in principle from the log envelope, there are a number of practical factors which make this unattractive.

The principal application identified to date for improved MLS angle processing in the presence of multipath is in the elevation angle estimation at low elevation angles such as arises during the flare operation. A simulation and (limited) experimental comparison was made between the SEP processor currently suggested for TRSB flare processing, several contemporary radar low elevation angle estimation techniques, and the extended Pisarenko/AR (ME) technique. It was found that the ME-based technique gave the best performance in two coherent signal simulations similar to those used for MLS flare assessment, but that performance differences were much less clear in experimental situations characterized by a multiplicity of reflected signals. Since the experimental results were affected significantly by the small number of aperture samples utilized, it is suggested that additional measurements be made with a more realistic number of elements.*

Improved angle processing in the presence of diffraction multipath has many of the same features as the elevation angle processing discussed in Section C except that it is not clear a priori which "side" of the direct signal will be corrupted by multipath. Since the aperture sampling techniques discussed here do not require a priori data, the simulation results of Section C are directly applicable. On the other hand, the operationally most significant diffraction problem identified to data - azimuth shadowing by taxiing aircraft - is most appropriately modeled by two or more coherent inbeam multipath signals which may lie on either side of the direct signal. Our preliminary investigations (recall Ch. V) suggest that the resolution performance of the various adaptive algorithms may be significantly degraded in such multiple-signal cases. More detailed simulation and experimental studies are needed here to see if a significant improvement can be obtained.

---

*The use of a COMPACT array in which each individual (overlapping) element has a significant pattern rolloff at the horizon [28] would be advantageous in terms of reducing the number of significant multipath signals.

VIII. SUMMARY AND CONCLUSIONS

In this report, we have examined the extent to which sophisticated signal processing techniques can improve the performance of navigation and/or ATC surveillance angle-determining systems. In this chapter we want to highlight our principal results and to indicate promising avenues for future work.

Our effort concentrated on improving multipath performance since

(1) theoretical performance bounds such as the Cramer-Rao bound discussed in Chapter II suggest that the likelihood of substantial improvement (e.g., greater than 6 dB in effective SNR) in performance against additive front end noise is unlikely,

(2) reducing errors due to instrumention effects would be very system/implementation dependent, and

(3) multipath represents an important factor in most contemporary ATC system designs as well as being a primary cause of out-of-tolerance performance.

However, both front-end noise and instrumention effects were considered, since it turns out that these error sources limit the capacity for improved multipath performance.

The analysis is chiefly concerned with improving the capability of antenna systems such as the FRG DAS that estimate angle of arrival. Such systems are utilized widely[*] and are theoretically equivalent to idealized versions of other systems (e.g., the TRSB subsystem of the MLS). Two approaches to obtain improved capability were considered:

(1) the spectral approach whereby one attempts to estimate the angular power spectrum of a stochastic process which models the received waveform at the antenna aperture. In this approach, the peaks in the (angle) spectrum are identified with the arrival angles of plane waves.

---

[*] e.g., The air traffic control radar beacon system (ATCRBS) and skin tracking radar. Also, some systems in development [e.g., the full capability Beacon Collision Avoidance System (BCAS)] will estimate angle of arrival in the presence of interference.

(2)   the parametric approach where plane-wave parameters are
       estimated explicitly (as in the ML probability method) or
       implicitly (as in the Prony method) from the signals
       received at the antenna elements.

These approaches are not mutually exclusive in the senses that some algorithms
(e.g., AR) may combine features of both approaches.

A.   Spatial Covariance and Consequential Results

     The central role of the signal covariance function was developed in
Chapter II. We introduced the notion of a "signal in space" covariance which
characterizes the extent that the received signals are correlated at various
times.   The subsequent analysis was concerned with two important limiting
cases:

   1.   Incoherent Signals

          The diagonal elements of $\underline{P}$ represent the power
       levels of the signals-in-space.   If the off-diagonal
       elements are identically zero, then the signals-in-space
       are said to be (jointly) uncorrelated and, for an ideal
       array, the signal covariance matrix $\underline{S}$ will exhibit the
       Toeplitz property; $S_{mn} = C_{m-n}$.   Any signal with a
       Toeplitz covariance matrix was said to be <u>spatially</u>
       stationary and the covariance matrix referred to as a
       <u>correlation</u> matrix.

          The stationary case occurs most naturally when the
       signals-in-space originate from (statistically) indepen-
       dent sources.   Examples of uncorrelated sources in ATC
       systems include temporarily overlapping signals from
       different transmitters (e.g., fruit or garble in DME and
       ATCRBS, DABS and BCAS) or multipath signals with a
       relative Doppler frequency which is high relative to the
       system measurement rate.

   2.   Coherent Signals

          If specular reflections or diffraction multipath is
       present, the signals-in-space are no longer uncorrelated
       since their relative phases and amplitudes are determined
       by geometry and the physical environment (terrain
       features, buildings, etc.).   When the signals-in-space
       maintain a fixed amplitude and phase relationship over an
       entire (time) observation interval, the signals-in-space

were said to be _perfectly_ correlated and $\underline{P}$ is a singular
matrix of the form $\underline{p} \ \underline{p}^H$.

By considering certain asymptotic algebraic properties of the signal
covariance estimates, it was possible to obtain results for the classical
problem: how many plane waves can be identified and estimated with a given
array? Specifically, we found that a uniform line array with L identical
elements can asymptotically[*] resolve:

   (1) as many as L-1 signals if the signals are incoherent,

   (2) as many as L/2 _coherent_ signals if subaperture spatial
       averaging of the type shown in Fig. 8-1a is utilized, and

   (3) two _coherent_ signals if only forward-backward smoothing
       (see Fig. 8-1b) is used and the relative phase is not
       near 0° or 180°.

A definitive statement of the combined effects of spatial smoothing and
forward/backward averaging in the coherent case remains elusive. For example,
it is probably true that forward/backward averaging increases the coherent
signal capability of an array "most" of the time, provided the number of
signals present does not exeed 2L/3. However, specific counterexamples are
easily constructed. For example, when the received signals are all inphase,
the theoretical capability of an array is not improved by forward/backward
averaging.

B. Probability Function Considerations

    Chapter III and Appendix E considered the use of maximum likelihood
(probability) estimates for the parameters (amplitude, phase, and arrival
angle) of several plane waves given the observed data. Both single-and
multiple-snapshot estimators were considered. These estimates involve
maximization of a nonlinear function of the received data over the angle-of-
arrival parameters. Numerical optimization of the function does not seem

_____

[*] i.e., as the number of time observations ("snapshots") approaches infinity.

8-3

ACTUAL APERTURE   OVERLAPPING SUBAPERTURES

AVERAGE OF SUBAPERTURE DATA STATISTICS TREATED
AS SINGLE APERTURE STATISTIC

(a)

S_1
S_2
S_3
S_4
S_5
S_6
S_7

S_7^*
S_6
S_5
S_4
S_3
S_2
S_1

S_i = RECEIVED SIGNAL
      AT i^{th} ELEMENT

ACTUAL
APERTURE

"EXCHANGED"
APERTURE

AVERAGE OF APERTURE DATA STATISTICS TREATED
AS SINGLE APERTURE STATISTIC
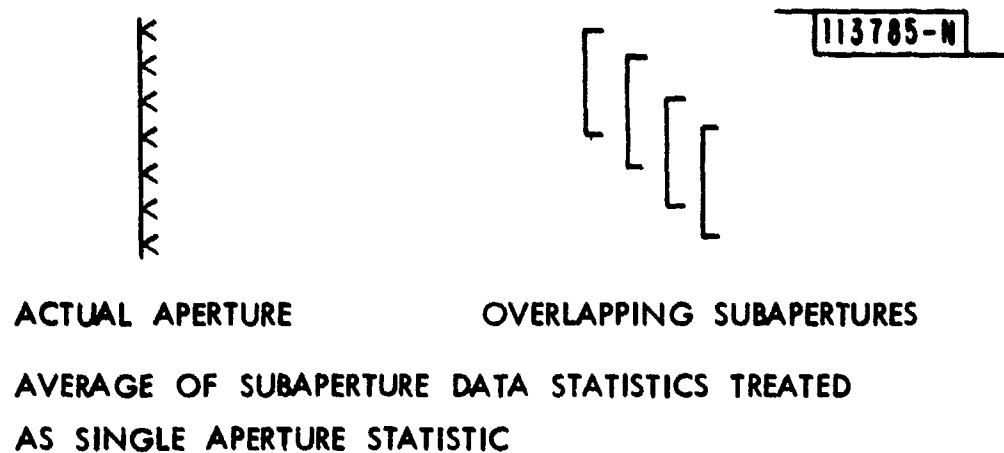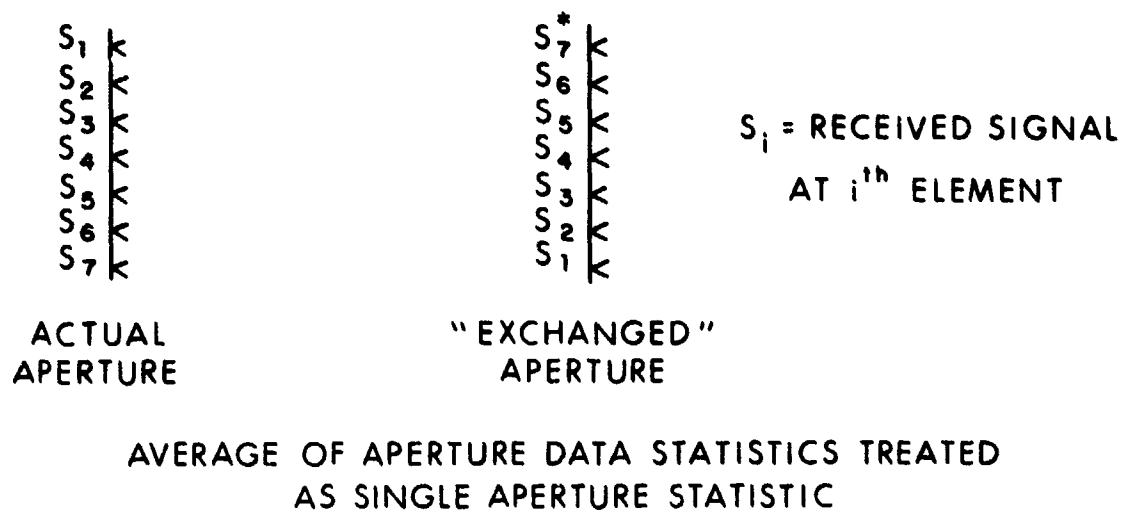
(b)

Fig. 8-1. Illustration of: (a) Subaperture Spatial Averaging, (b) Forward-
Backward Spatial Smoothing.

practical for real-time applications and hence was not investigated in depth during the course of this study.

However, several intermediate results are viewed as noteworthy:

(1) a gradient projection technique for recursively computing the angle parameter estimates was shown to depend on the Fisher information matrix which arises in Cramer-Rao (CR) bounds on the ML performance.

(2) the estimator for multiple observations of coherent signals is shown to involve the sample covariance matrix which was considered in Chapter II. A suboptimal approximation to this result is seen to involve use of the principal component (i.e., $\lambda_1 \, e_1 \, e_1^H$ where λ_1 and e_1 the largest eigenvalue and corresponding eigenvector) of the covariance matrix as if the principal component was a single array observation.

Result (2) is of interest because it relates the seemingly disparate concepts of covariance and covariance eigenanalysis to probabilistic considerations.

Explicit Cramer-Rao bounds on the angle-of-arrival standard deviation were obtained for the case of a single time observation of one or two plane waves in additive noise. These results extend those obtained by other investigators [15, 16]. One important contribution is that the degradation in effective signal-to-noise ratio when two signals are present can be written as a product

$$\eta \, (\Delta\alpha) \, [1 - |f(\Delta\alpha)|^2 \cos^2 \Delta\beta \,]$$

where the array efficiency factor, η and the array coherency factor f depend only on the angular separation $\Delta\alpha$ between the plane waves and $\Delta\beta$ depends on the relative rf phase. This representation clarifies the rather large performance differences which occur for inphase/out-of-phase signals as opposed to quadrature phase.

On the other hand, several apparently paradoxical features of the two signal CR bound were not resolved. These include:

(1) no dependence on relative signal amplitudes in the expression for angle-of-arrival variance. This seems odd in that it suggests a small amplitude signal very near a much larger signal may cause a severe error in the angle-of-arrival estimate of the desired (larger) signal, and

(2) a very rapid increase in the bound on angle variance as
    the separation angle decreases for 0°and 180° relative
    phase.   By contrast, practical angle-of-arrival
    estimators typically show a decreasing rms error as the
    separation angles decreased at 0° relative phase.

We have highlighted these paradoxical features since the CR bounds for most
radar estimation problems (see e.g., [34]) do not have such anomalies. It may
be that the restriction to unbiased estimator CR bounds generates some of
these paradoxical features and the use of alternative bounding techniques
[82, 83] may be useful.   On the other hand, the analysis in Appendix E
suggests that at typical SNR and small separation angles, a "realistic"
optimal processor which has no a priori knowledge of the correct number of
plane waves will simply fail to resolve the two signals.   Hence, the
"anomalous" behavior of the CR performance bound discussed above may not be
important in practical situations.

C.   Practical High Resolution Techniques

Next we considered a number of more practical estimation techniques. The
extended Prony/autoregressive (AR) technique was developed both as an
extension of the classical Prony method[*] and from the viewpoint of fitting
an-all pole angular spectrum model to the received data. This method was seen
to be superior in terms of resolution and angle estimation performance to
other common maximum entropy approaches (specifically the Burg and
autocorrelation methods) for the case of a small number of time samples and
equivalent to these in the asymptotic limit for incoherent signals.

The spatially averaged covariance estimates which arise out of this
technique were found to have useful application to two other high-resolution
techniques:

(1) the maximum likelihood method (MLM) due to Capon, and

(2) covariance eigenanalysis techniques (e.g., MUSIC)

---

[*]Mathematically equivalent to an approach suggested at one time for DLS
processing.

which hitherto had been found to yield very poor performance when the signals were coherent. This is an important result since it was previously believed (see, e.g., [1, 13]) that the MLM technique could not be used successfully in situations with coherent signals.

The covariance eigenanalysis techniques appear particularly attractive in cases where there are a large number of time observations. In the limit of infinite time observations, these techniques permit one to effectively eliminate the contribution of additive noise to the covariance. When the noise contribution is eliminated, it is then fairly easy (e.g., by the Pisarenko method) to exactly determine the arrival angles of the various plane waves. However, in the more common case of a small number (e.g., 1-10) of time observations, the noise-reduction advantages of eigenanalysis techniques are not always evident.

ARMA techniques were also investigated because sum of sinusoids (plane waves) in noise satisfies a difference equation similar to the equation describing an ARMA process. On the other hand, the presence of zeros (as well as poles) in the assumed data model significantly complicates the parameter estimation problem, and we were not able to develop an effective ARMA algorithm within the time frame of the present study.

D. Performance Assessment of Various Techniques

At the time this work commenced, the principal performance assessment of improved techniques had consisted primarily of simulation studies on:

(i) incoherent signals with the observable being the ensemble correlation matrix, and/or

(ii) coherent signals with a single time observation wherein the published result was a angular spectrum plot.

The ensemble correlation matrix results can provide reasonable bounds on the multiple-snapshot/incoherent signal performance of several algorithms (e.g., MLM and AR/ARM), but they tend to give overly optimistic results for certain others (especially MUSIC, ARMA, and EAR). The single-snapshot angular spectra

3-7

displays suffer from the lack of precise quantitative criteria as to the improvement, especially since visual criteria such as sharpness of peaks are not necessarily a good measure of resolution or angular accuracy.

To remedy these deficiencies, a Monte Carlo simulation based on statistical decision theory was developed in which the angular spectra (or, model parameters) were analyzed to provide plane-wave "target reports" (i.e., angle-of-arrival and power estimates). These target reports are then assigned to the actual plane waves so that statistical decision theoretic error measures (e.g., false alarm and miss probability, rms and bias error) can be determined. By setting (application dependent) quantitative bounds on each error measure, one can then achieve a quantitative expression of successful performance.

The Monte Carlo simulations and ensemble covariance studies showed that certain of the advanced signal processing techniques (especially, the ARM and eigenanalysis approaches) have the potential for significant (e.g., factors of 3 to 10) improvement in performance for reasonable SNR (e.g., 30):

1. For incoherent signals or, quadrature-phase coherent signals, two plane waves can be successfully resolved[*] at separation angles as small as 0.1 BW with 10 or more time observations. At a lower array SNR (e.g., 20 dB) the minimum angular separation is on the order of 0.3 BW.

2. Two inphase (e.g., relative phases near 0° or 180°) coherent plane waves with 10 or more time observations can be successfully resolved at separation angles down to 0.5-0.7 BW. The degradation with SNR is relatively slow (e.g., 20 dB array SNR yields a minimum resolvable angular separation of 0.8 to 1.0 BW, 40 dB SNR yields a minimum angular separation of 0.2-0.35 BW). On the other hand, if only a single time observation is available the required SNR for resolution will increase significantly (e.g., over 15 dB at 0.5 BW separation angle).

---

[*]Successful resolution here is defined as miss and false-alarm probabilities of less than 10% with angle-of-arrival rms and bias errors on all received plane waves of less than 0.05 BW.

Conventional angle determination techniques (e.g., monopulse or the TRSB thresholding processor) will typically fail these resolution criteria for separation angles below 1.5-2 beamwidths.[**] Thus, the above performance results represent a factor of 5 to 15 improvement in minimum separation angle in the incoherent or quadrature-phase cases and a factor of 2-3 improvement in the coherent inphase case.

The independent Gaussian additive noise considered in the Monte Carlo and ensemble covariance simulation is one of several error sources for an angular direction determining system. Other error sources (e.g., receiver channel drifts, A/D converter errors, array collimation errors) were examined primarily in the context of experimental measurements. The measurements in a controlled environment at the Lincoln Laboratory antenna range generally agreed well with the Monte Carlo results. For example, the angle error using a single time observation to estimate the larger of two coherent plane waves (ρ = 0.7) at an array SNR of 33 dB was:

(i) < 0.07 BW for all relative phases which were 30° removed from 0° or 180°, and

(ii) < 0.07 BW for all relative phases at a separation angle of 1 BW.

At 0.5 BW separation angle, the weaker of the two plane waves could be resolved and located to within 0.1 BW for all relative phases which were approximately 45° removed from the 0°/180° phase conditions.

The (limited) simulation studies of the effects of differential amplitude and/or phase errors in the measured array data indicate that the performance for coherent inphase signals degrades sharply in terms of detection probability and angular errors when the amplitude errors exceed 1 dB and phase errors exceed 3°. The sensitivity to errors with quadrature-phase coherent signals was much less pronounced. Much more extensive studies in this area are needed to clarify the performance degradation as a function of SNR, signal environment, and number of elements.

[**]Depending on the relative signal amplitudes.

Several examples using the high-resolution spectrum estimation techniques to analyze terrain reflection data were presented. These examples illustrate the problems which arise in attempting to identify the nature of the received signals. Specifically, with the techniques studied to date, it has been quite difficult to distinguish between an environment consisting of discrete plane waves with low SNR and small separation angles and a "mixed" environment with discrete plane waves as well as 'diffuse' signals from an extended angular region. ARMA estimators may offer some utility in such cases, particularly since the environment analysis application often does not necessitate real-time operation.

E. Hardware Implementation of Promising Techniques

Key hardware implementation issues for a sampled-aperture receiving system were identified based on several years of experience with an L-band/C-band system and via simulation studies. These investigations have identified amplitude and phase equalization among the various receiver channels* as a principal problem. In particular, coherent inphase signal performance degrades significantly with amplitude errors greater than 1 dB and phase errors above 3°. Based on our experience, it appears that a combination of careful hardware design and real-time calibration should be capable of achieving the desired tolerances provided that the array element patterns do not sufer significant differential variations with time and/or temperature.

Signal quantization and round-off errors, which had previously been important factors in real-time implementations, currently appear to be much less of a problem due to the significant progress in A/D converter capability and the ready availability of high-speed, low-cost multiplier chips. The spectrum analysis computation costs could be significant since a very fine (angle) grid may be required. However, one of the most promising techniques (ARM) requires only rooting of a polynomial. Simulation experiments suggest

---

*For an air-derived processor (e.g., the TRSB receiver), the corresponding requirement is for collimation of the paths from the ground-transmitter local oscillator (or, transmitter) to the individual elements.

that the root-finding operation will require significantly less computer time than angular spectrum analysis for arrays with a small to moderate number of elements.

F.   Application of the Techniques to TRSB Signal Processing

The TRSB signal from a line array is the Fourier transform of the wavefront which would be received by a ground-based receiving system with the same antenna aperture. Thus, a (relatively short) investigation was made of the possibility for implementing some of the more promising signal processing techniques in a TRSB receiver. The principal problem identified was the loss of phase information in the TRSB envelope and two approaches were considered:

(a)   recovery of phase information using the DPSK function preamble, and

(b)   wavefront reconstruction directly using the fact that phase and magnitude are closely related for minimum-phase signals.

The phase information recovery via DPSK demodulation requires additional study to understand how frequency offsets between the demodution frequency and TRSB carrier (due to low SNR and/or multipath scalloping) will degrade performance. Some of the difficulty here arises from the fact that the TRSB signal format scans at an uniform angular rate rather than uniformly in sine space.   Other difficulties include the lack of information at the TRSB receiver regarding the ground-system aperture (e.g., number of elements, spacing) and the type of aperture weighting used to reduce the antenna sidelobes.

The direct wavefront reconstruction is appealing theoretically. However, the algorithms which are currently available typically involve a numerically difficult gradient search which would not be appropriate for airborne receivers. Hence, this technique does not seem practical.

An alternative signal processing approach which might be useful is to investigate high-resolution algorithms which use as their input the received signals from a beamformer array (e.g., as from a Butler matrix) since the TRSB

8-11

complex envelope sampled at suitable points in time would closely approximate the output of such a beamformer array.

To provide quantitative insight into the TRSB performance improvement which might be achieved with sophisticated signal processing techniques, simulation and field data comparisons were made between several low angle elevation angle estimators that might be used for flare guidance. Specific comparison was made between the TRSB single edge processor used for flare guidance, a AR-based estimator, and several contemporary low angle monopulse trackers. The AR-based estimator showed significantly better performance in the simulations of performance over flat terrain; however, the results with field data from irregular terrain showed no clear advantage for any of the techniques. Additional measurements using a more representative receiving array* and aircraft geometry are needed here.

G.   Concluding Remarks

Our principal conclusions can be briefly summarized as follows:

(1) There seems little likelihood of significantly improved angle-of-arrival estimation performance for cases where only one desired signal is present and signal-to-noise ratio is the principal limitation.

(2) Improvements of 2-3 in minimum resolvable angular separation are possible against coherent interference (e.g., multipath) with high SNR (e.g., 30 dB) and an appropriate array geometry which permits spatial smoothing.

(3) Order-of-magnitude improvements in minimum resolvable angular separation are achievable against incoherent interference with moderate SNR (20 dB) and modest time averaging. The most obvious case where such incoherent interference arises in current ATC systems is synchronous garble in ATCRBS/BCAS.

(4) Computation hardware for certain of the high-performance algorithms seems readily achievable at low to moderate cost.

---

*Horizon limitations (due to trees) necessitated the use of arrays with a small number of elements in order to obtain an inbeam multipath condition for the field data.

8-12

(5) Receiver calibration/monitoring will be a potentially
    significant factor in practical realization. The re-
    quired amplitude and phase accuracies need additional
    study.

The preceding portions of this chapter have identified a number of areas
which require additional work. The most pressing items are as follows:

(1) improved performance bounding techniques so that perfor-
    mance benchmarks can be defined.

(2) continued work on estimators which offer considerable
    promise [e.g., EAR because of its good detection perfor-
    mance, ARMA for applications in which the shape of the
    angular power spectrum (as opposed to peak locations is
    important)], but have undesirable features in their
    current versions. Assessment of nonequally spaced line
    arrays and/or two-dimensional arrays is also needed since
    a number of current (e.g., VOR) or proposed (e.g.,BCAS
    and DAS) systems use such arrays.

(3) inclusion of aperture sample phase and amplitude errors
    into the decision theoretic (Monte Carlo) simulation so
    that the required channel equalization, calibration, and
    monitoring can be more precisely quantified.

(4) detailed investigation of specific candidate systems to
    determine the total system implication of using high-
    resolution signal processing techniques vis-a-vis current
    methods.

References

1. W. Gabriel, "Spectral analysis and adaptive array superresolution techniques", Proc. IEEE 68, 654-666 (1980).

2. D. Childers (editor), Modern Spectrum Analysis (IEEE Press, New York 1978).

3. H. Cox, "Resolving Power and Sensitivity to Mismatch of Optimum Array Processors," J. Acoust. Soc. Amer. 54, No. 3, 771-785 (1973).

4. W. I. Wells, "Verification of DABS Sensor Surveillance Performance (ATCRBS mode) at Typical ASR Sites Throughout CONUS", Project Report ATC-79, Lincoln Laboratory, M.I.T. (10 Nov. 1977), DDC AD-A05112817.

5. J. Evans, D. Sun, D. Shnidman, and S. Dolinar, "MLS Multipath Studies Phase 3 Final Report, vol. 1: Overview and Propagation Model Validation/Refinement Studies", Project Report ATC-88, Lincoln Laboratory, M.I.T. (Apr. 1979) FAA-RD-79-21, DTIC AD-A08782712.

6. J. Capon, "Multipath Parameter Computations for the MLS Simulation Computer Program", Project Report ATC-68, Lincoln Laboratory, M.I.T. (8 April 1976) FAA-RD-76-55, DDC AD-A024350/10.

7. "Development of a Discrete Address Beacon System," Quarterly Technical Summary, 1 April - 30 June 1976, Lincoln Laboratory, M.I.T. (1976).

8. "Development of a Discrete Address Beacon System," Quarterly Technical Summary, 1 October - 31 December 1976, Lincoln Laboratory, M.I.T. (1977), DDC AD-A-37130/2.

9. "The DME-Based-Landing System, DLS," as proposed by the Federal Republic of Germany developed by Standard Elektrik Lorenz AG. and Siemens, AG., Sept. 1975

10. M. Bohm, "A Simple Multipath Error Reduction Technique For Single Site DF Systems", AGARD Conf. Proc. No. 209 on Propagation Limitations of Navigation and Positioning Systems, London, 1976.

11. H. Van Trees, Detection, Estimation and Modulation Theory Part I (Wiley, New York, 1968).

12. J. Capon, "High-Resolution Frequency-wave Number Spectrum Analysis", Proc. of IEEE 57,1408 (1969).

13. W. D. White, "Angular spectra in radar applications," IEEE Trans. Aerosp. Electron Syst. AES-15, 895-899 (1979).

14. D. Barton, Radar Resolution and Multipath Effects (Dedham, Mass, Artech

Press, 1975).

15. J. R. Sklar and F. C. Schweppe, "The Angular Resolution of Multiple Targets," Proc. of IEEE.

16. G. E. Pollon, "On the Angular Resolution of Multipath Targets," IEEE Trans. Aerospace Electron Systems (Corresp.), AES-3, 145-148 (1967).

17. H. Akaike, "Statistical Predictor Identification," Ann. Institute Statistical Mathematics 22,205 (1970).

18. W. White, "Low Angle Radar Tracking in the Presence of Multipath," IEEE Trans. Aerospace Electron Systems AES-10, 838-852 (1974).

19. R. T. Lacoss, "Autoregressive and Maximum Likelihood Spectral Analysis Techniques, " in Aspects of Signal Processing, pt. 2, G. Tucconi, Ed. (Reidel, Boston,1977).

20. J. Makhoul, "Linear Prediction: A Tutorial Review," Proc. IEEE 63, 561-580 (1975).

21. M. Kaveh, "High Resolution Spectral Estimation for Noisy Signals,"IEEE Trans. Acoust., Speech, and Signal Processing ASSP-27, 286-287 (1979).

22. G.E.P. Box and G. M. Jenkins, Time Series, Forecasting and Control (Holden Day, San Francisco, CA , 1970).

23. S. Bruzzone and M. Kaveh, "On Some Suboptimum ARMA Spectral Estimators," IEEE Trans. Acoust., Speech, and Signal Processing ASSP-28, 753-755 (1980).

24. International Mathematics and Scientific Program Library (1978).

25. J. Cadzow, "High Performance Spectrum Estimation - a New ARMA Method," IEEE Trans. Acoust., Speech, and Signal Processing ASSP-28, 524-529 (1980).

26. S. Kay, "A New ARMA Spectral Estimator," IEEE Trans. Acoust., Speech, and Signal Processing ASSP-28, 585-588 (1980).

27. W.B. Kendall, "Unambiguous Accuracy of an Interferometric Angle-Measuring System," IEEE Trans. Space Electron and Telemetry 11,62-70 (1965).

28. J. Evans, S. Dolinar, D. Shnidman, and R. Burchsted, " MLS Multipath Studies Phase 3 Final Report Volume II: "Development and Validation of Model for MLS Techniques," Project Report ATC-88, Vol. II, Lincoln Laboratory, M.I.T. (7 February 1980), DTIC AD-A08001/3.

29. J. Evans and S. Dolinar, "MLS Multipath Studies Phase 3 Final Report Volume III: "Comparative Assessment Results," Project Report ATC 88, Vol. III, Lincoln Laboratory, M.I.T. (in press).

30. R. Schmidt, "Multiple Emitter Location and Signal Parameter Estimation," Proc. RADC Spectrum Estimation Workshop, Rome, N.Y., 1979.

31. S. Lang and J. McClellan, "Frequency Estimation with Maximum Entropy Spectral Estimators," IEEE Trans. Acoust., Speech, and Signal Processing ASSP-28, 716-724 (1980).

32. S. Lang, "Near Optimal Frequency/Angle of Arrival Estimates Based on Maximum Entropy Spectral Techniques," in Proc. 1980 IEEE Intl. Conf. on Acoustics, Speech and Signal Processing, 829-832 (1980).

33. K. Kohler, "Synthetic Radio Direction Defining Methods with Virtual Antenna Patterns",Electrical Comm. 48,299-304 (1973)

34. D. K. Barton and H. R. Ward, Handbook of Radar Measurement (Prentice-Hall, New Jersey,1969).

35. U. Nickel, "Detection and Parameter Estimation of Closely Spaced Multiple Targets," IEEE Symp. on Info. Theory, Ithaca, N.Y., 1977.

36. H. Birgenheier, "Parameter Estimation of Multiple Signals," Engineering System Science Ph.D. Dissertation, Univ. California at Los Angeles, 1972.

37. I. Kupiec, "Experimental Verification of the Performance of the Aperture Sampling Technique," Technical Note 1975-45, Lincoln Laboratory, M.I.T., (15 September 1975), DDC AD-A016779/1.

38. I. Kupiec, "Compensation of Multipath Angular Tracking Errors in Radar," Technical Report 501, Lincoln Laboratory, M.I.T. (20 March 1974), DDC AD-781100/3.

39. J.E. Howard, R.A. Birgenheier, and G.O. Young, "Application of Learning Techniques to AMTI Radar Problems," Technical Report AFAL-TR-72-240 (August 1972).

40. T. Ulrych and M. Ooe, "Autoregressive and Mixed Autoregressive - Moving Average Models and Spectra," in Nonlinear Methods of Spectral Analysis, S. Haykin, Ed. (Springer Verlag, N.Y., 1979).

41. M. Morf, B. Dickinson, T. Kailath, and A. Viera, "Efficient Solution of Covariance Equations for Linear Prediction," IEEE Trans. Acoust, Speech, and Signal Processing ASSP-25, 429-433 (1977).

42. J. E. Evans, D. Karp, R. LaFrey, R. J. McAulay, and I. G. Stiglitz, "Experimental Validation of PALM - A System for Precise Aircraft Location," Technical Note 1975-29, Lincoln Laboratory, M.I.T. (29 April 1975), DDC AD-A010112/1.

43. Federal Republic of Germany, "The Need for Integrated Navigation System in the TMA," ICAO All Weather Operation Division Working Paper AWO/78-WP/56, (1978).

44. J. Evans, "L-band DME Multipath Environment in the MLS Approach and Landing Region," Project Report ATC-116, Lincoln Laboratory, M.I.T. (March 1982), FAA-RD-82/19.

45. Weiss, L. and R. N. McDonough, "Prony's Method, Z-Transforms, and Pade Approximation," SIAM Rev. $\underline{5}$, 145 (1963).

46. Van Blaricum and Mittra, "Problems and Solutions Associated with Prony's Method for Processing Transient Data," IEEE Trans. Antennas Propag $\underline{AP-26}$, 174-182 (1978).

47. F. B. Hildebrand, Introduction to Numerical Analysis (McGraw-Hill, New York, 1956).

48. R. Prony," Essai experimental et analytique, etc.," Paris, J. l'Ecole Polytechnique $\underline{1}$, cahier 2, 24-75 (1975).

49. O. L. Frost, "An Algorithm for Linearly Constrained Adaptive Array Processing," IEEE Proc. $\underline{60}$, 926-935 (1972).

50. A. J. Berni, "Angle of Arrival Estimation Using an Adaptive Antenna Array," IEEE Trans. Aerospace Electron Systems $\underline{AES-11}$, 278-284 (1975).

51. L. Marple, "High-Resolution Autoregressive Spectrum Analysis Using Noise Power Cancellation," IEEE International Conf. on Acoustics, Speech, and Signal Processing, Tulsa, Oklahoma, 10-12 April 1976.

52. R. Nitzberg, "Application of Maximum Likelihood Estimation of Persymmetric Covariance Matrices to Adaptive Processing," IEEE Trans. Aerospace Electron. Systems $\underline{AES-16}$, 124-127 (1980).

53. Y. L. Chu, "Optimum Complex ARMA in Cascade Form," S.B. Thesis, M.I..T. Electrical Engineering Dept., May 1981.

54. R. Manasse, "An Analysis of Angular Accuracies from Radar Measurements," Group Report 32-34, Lincoln Laboratory, M.I.T. (6 Dec. 1955), DDC236166.

55. M. Marcus, "Basic Theories in Matrix Theory," National Bureau of Standards, Applied Math Series, Vol. 57, (27 Jan. 1960).

56. B. Cantrell, W. Gordon, and G. Trunk, "Maximum Likelihood Elevation Angle

Estimates of Radar Targets Using Subapertures," IEEE Trans. Aerospace Electron. Systems <u>AES-17</u>, 213-220 (1981).

57. J. E. Howard, "A Low Angle Tracking System for Fire Control Radars," IEEE 1975 Int. Radar Conf. Rec., 412-417 (1975).

58. D. C. Rife and R. R. Boorstyn, "Single-Tone Parameter Estimation from Discrete-Time Observations," IEEE Trans. Inf. Theory <u>IT-20</u>, 591-598 (1974).

59. G. Trunk, B. Cantrell, and W. Gordon, "Probability Density of the Maximum Likelihood Elevation Estimation of Radar Targets," IEEE Trans. Aerospace Electron. Systems <u>AES-15</u>, 288-289 (1979).

60. B. Gold and C. Rader, <u>Digital Processing of Signals</u> (McGraw-Hill, N.Y. 1969), Chapter 8.

61. J. P. Burg, "A New Analysis Technique for Time Series Data," <u>NATO Advanced Study Institute on Signal Processing with Emphasis on Underwater Acoustics</u>, August 12-23, 1968.

62. J. P. Burg, "Maximum Entropy Spectral Analysis," PhD Thesis, Stanford Univ. May 1975.

63. N. O. Anderson, "On the Calculation of Filter Coefficients for Maximum Entropy Spectral Analysis," Geophysics <u>39</u>, 69-72 (1974).

64. A. B. Baggeroer, "Confidence Intervals for Regression (MEM) spectral Estimates," Tech. Note 1975-41 Lincoln Laboratory, MIT, (18 June 1975), DDC AD-A-012286/1; and IEEE Trans. Inf. Theory <u>IT-22</u>, 534-545 (1976).

65. T. J. Ulrych and R. W. Clayton, "Time Series Modeling and Maximum Entropy," Phys. Earth Plan. Int. <u>12</u>,188-200 (1976).

66. A. H. Nuttal, "Spectral Analysis of a Univariate Process with Bad Data Points, via Maximum Entropy and Linear Predictive Techniques," Naval Underwater Systems Center, New London, CT, Tech. Rep. 5303 (Mar. 26, 1976).

67. V.F. Pisarenko, "The Retrieval of Harmonics from a Covariance Function," Geophys. Royal Astron. Soc. <u>33</u>,347-366 (1973).

68. U. Grenander and G. Szego, <u>Toeplitz Forms and Their Application</u> (Univ. of California Press, Berkeley, 1958).

69. S. S. Reddi, "Multiple Source Location - A Digital Approach," IEEE Trans. Aerospace Electron Systems <u>AES-15</u>,95-105 (1979).

70. "Development of a Discrete Address Beacon System," Quarterly Technical Summary, 1 Oct. - 31 Dec. 1973, Lincoln Laboratory, M.I.T. (1 January 1974), FAA-RD-74-8, DDC AD-776118/2.

71. J. L. Allen, "The Theory of Array Antennas," Lincoln Laboratory Technical Report No. 323, (25 July 1963).

72. N. R. Goodman, "Statistical Analysis Based on Certain Multivariate Complex Gaussian Distribution," Ann. Math Stat. 34,152-177 (1963).

73. F. J. Harris, "On the Use of Windows for Harmonic Analysis with the Discrete Fourier Transform," Proc. IEEE 66,51-83 (1978).

74. T. F. Quatieri and A. V. Oppenheim, "Iterative Techniques for Minimum Phase Signal Reconstruction from Phase or Magnitude," Technical Note 1980-34, Lincoln Laboratory, M.I.T. (1 August 1980), DTIC AD-A091111.

75. J. R. Fienup, "Reconstruction of an Object from the Modulus of its Fourier Transform," Opt. Lett. 3, 27 (1978).

76. A. V. Oppenheim and R. W. Schafer, Digital Signal Processing (Prentice-Hall, Englewood Cliffs, New Jersey, 1975).

77. B. Gold, A. V. Oppenheim, and C. M. Rader, "Theory and Implementation of the Discrete Hilbert Transform," Proceedings of the Symposium on Computer Processing in Communication, Vol. 19 (Polytechnic Press, New York, 1970).

78. P. Dax, "Keep Track of that Low-Flying Attack," Microwaves 36-53 (1976)."

79. J. Beneke, D. Wightman, A. Offt, and C. Vallone, "TRSB Multimode Digital Processor," FAA-RD-78-84 (April 1978).

80. J. Warren, "Flight Tests of the Microwave Landing System Multimode Digital Processor," FAA Technical Center Report No. FAA-CT-80-19 (Jan. 1981).

81. H. Kahn, "Applications of Monte Carlo," RAND Corp. Research Memorandum RM-1237-AEC (April, 1956).

82. M. Wax and J. Ziv, "Improved Bounds on the Local Mean-Square Error and the Bias of Parameter Estimators," IEEE Trans. Inf. Theory IT-XX, 529-530 (1977).

83. L. P. Siedman, Performance Limitation and Error Calculations for Parameter Estimation," Proc. IEEE, 38, 646-652 (1970)

84. ITT/Gilfillan, "Microwave Landing System (MLS) Development Plan as Proposed by ITT/Gilfillan During the Technique Analysis and Contract Definition Phase," FAA-RD-74-118 (Sept. 1972).

85. R. Kelly and E. LaBerge, "MLS Flare Low Elevation Angle Guidance

Considerations," National Telecomm. Conf. Record, volume 1, Houston, Texas, Nov. 1980.

86. C. Wightman, "Flare Pole Tests of the MLS Multimode Digital Processor," CALSPAN Corp., TN-15, June 1978.

87. S. Haykin and J. Reilly, "Mixed Autoregressive - Moving Averaging Modeling of the Response of a Linear Array Antenna to Incident Plane values," Proc. IEEE 68, 622-623 (1980).

88. A. Cantoni and L. Godara, "Resolving the Directions of Sources in a Correlated Field Incident on an Array," J. Acoust. Soc. Am. 64. 1247-1255 (1980).

89. D. Sun, Lincoln Laboratory, M.I.T., Project Report ATC-107 (in preparation).

90. R. Gallagher, Information Theory and Reliable Communication (Wiley, New York, 1968).

91. U. Grenander and G. Szego, Toeplitz Forms and Their Applications (Univ. of Calif. Press, Berkeley, 1958).

92. D. Luenberger, Optimization by Vector Space Methods (Wiley, New York, 1969).

93. J.Edward and M. Fitelson.,"Notes on maximum entropy processing," IEEE Trans. Inf. Theory. IT-17, 232-234 (1973).

94. A. Van den Bos, "An Alternative Interpretation of Maximum Entropy Spectral Analysis," IEEE Trans. Inf. Theory (Corresp.) IT-17, 493-494 (1971).

95. R. Deutsch, Estimation Theory (Prentice Hall, Englewood Cliffs, N.J., 1965).

96. J. Durbin, "The Fitting of Time-Series Models," Rev. Intern Statist Inst. 28,233-243 (1960).

97. N. Levinson, A Heuristic Exposition of Wiener's Mathematical Theory of Prediction and Filtering," J. Math. Phys. 25,110-119 (1947).

98. R.C. Buck, Advanced Calculus (McGraw-Hill, New York, 1956).

99. R. Bartle, The Elements of Real Analysis (Wiley, New York, 1964).

APPENDIX A

C-BAND ANTENNA ARRAY

The C-band antenna array, as shown in Fig. A-1,[*] consists of 30 antenna elements with two different element spacings. The elements in the upper half of the array have 3.1488 wavelengths (λ) spacing (distance between the centers of two adjacent elements) and those in the lower half of the array have 1.5744λ spacing. In the actual C-band measurement system, only the lower 29 elements are used.[**] This yields the usable array aperture of 56.7λ.

The individual antenna element consists of an E-plane sectoral horn and a twisted waveguide section with the coaxial feed at the far end of the wave-guide section. This twisted waveguide section was necessitated by the desired broad-wall to broad-wall waveguide coupling for the RF test signal injection for the RF calibration as discussed in Chapter VI Section B-2. These 30 horn antennas are then attached to a directional coupler assembly which is a 12-foot-long 2" x 1" waveguide with specially designed coupling slots for inject-ing the RF test signal, as shown in Fig. 6-7.

The coupling (or isolation) between two adjacent horns was thoroughly investigated. The isolation between two horns was measured for the two most popular arrangements of two horns, as shown in Fig. A-2, i.e., E-plane to E-plane and H-plane to H-plane. Although the arrangement of 30 horns in our C-band array was E-plane to E-plane, the isolation measurement result for the H-plane to H-plane arrangement would be helpful, in case we had some misalign-ment of these horns along the array axis. The measurement results are shown in Fig. A-2. We see that the isolation for the H-plane to H-plane arrangement is very good, well above 40 dB for any spacing. However, the isolation for

---

[*]In this picture, the array radome was removed to show the actual antenna elements.

[**]The use of only 29 elements (instead of available 30) is to make each of the four RF front end shared by the same number of antenna elements, with the bottom element used as a phase reference channel which was connected to the reference RF/IF receiver channel as shown in figure 6-2(b).
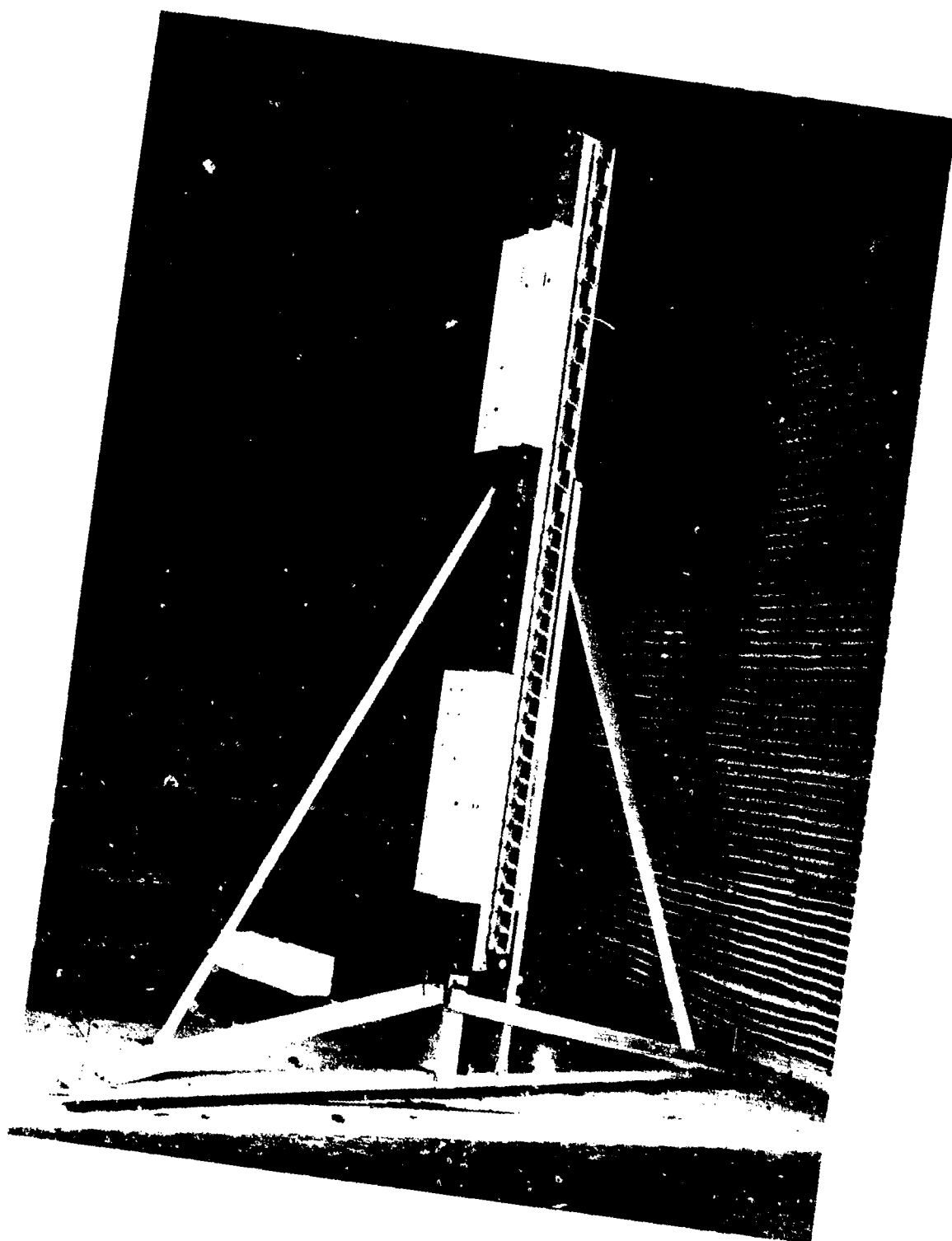
Fig. A-1. C-band array used for experimental studies.
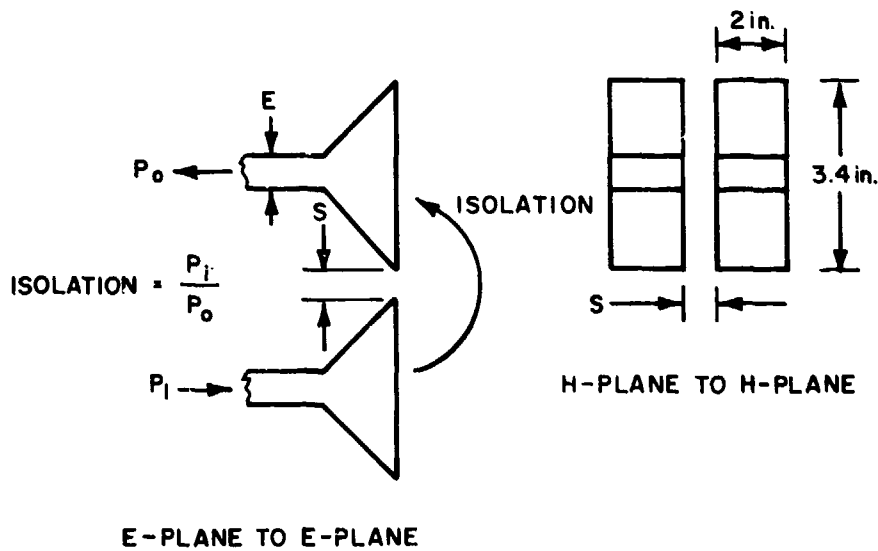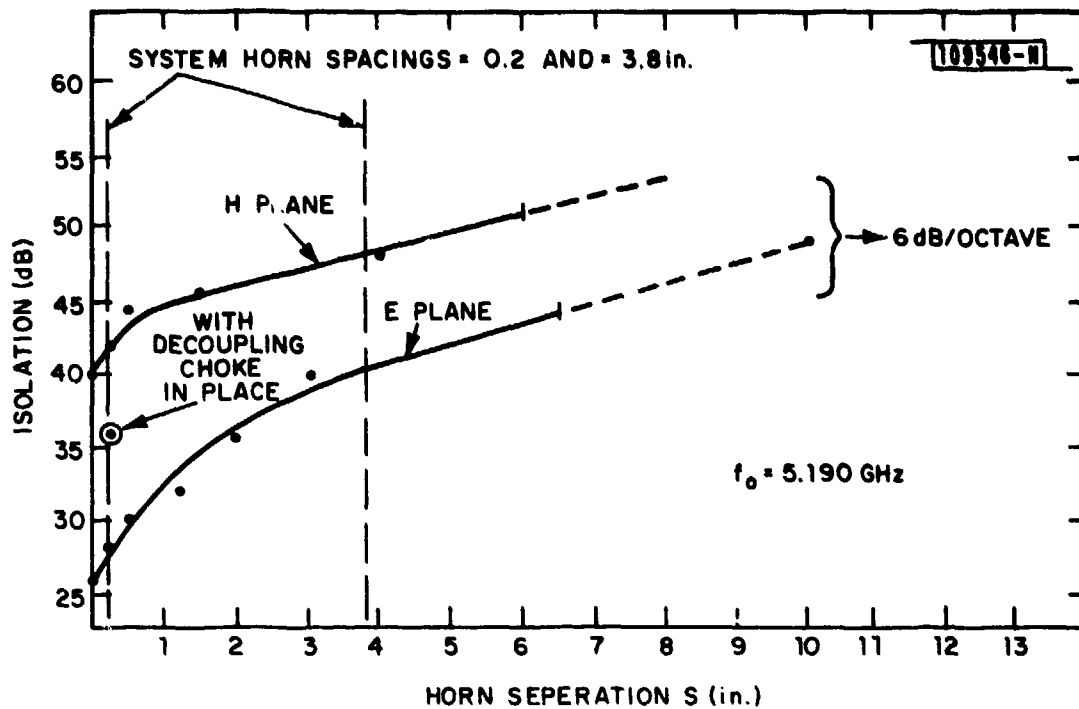
Fig. A-2. Horn-to-horn isolation versus spacing for two E-plane sectoral horns.

the E-plane to E-plane arrangement can be as low as 25 dB for very close spacing. As indicated in Fig. A-2, the isolation for the large element spacing in our C-band array is around 40 dB while that for the smaller spacing is only about 28 dB which is not good. To make the isolation between any two adjacent horns in our C-band array be at least better than 35 dB, an RF decoupling choke was installed in each space between two adjacent horns for the lower half of the array. With the choke in place the isolation between two adjacent horns with smaller spacing is improved from 28 to 36 dB, as indicated by the symbol in Fig. A-2. To further enhance the horn-to-horn isolation, when the antenna array was installed inside the radome, the distance between the array horn surface to the radome was adjusted to produce a 50-dB horn-to-horn isolation, as shown in Fig. A-3.

The connection between the 29 antenna elements and the 5 RF front ends is made through 5 SP10T RF switches as shown in Fig. 6-2(b). The bottom antenna element is used as phase reference and connected through one SP10T* switch to the reference RF front end. The other 28 antenna elements are equally divided among 4 SP10T switches. So, in the measurement, every 7 antenna elements are time multiplexed through one SP10T switch to share one RF front end. Since there are 10 input ports in one SP10T switch, 7 input ports with the best RF isolation and the least loss are chosen to be connected to 7 antenna elements. The measurements indicate that the switch port RF isolation is better than 40 dB for all switches used.

The design dimensions for the E-plane sectoral horn are given in Fig. A-4. Since the radiation characteristics (such as directivity, 3-dB beamwidth, sidelobe level, etc.) of a horn antenna can be fairly accurately determined from the horn dimensions, no radiation pattern measurements were made for the 29-horn antennas used in our C-band array. The calculated values are given below:

---

*The use of a SP10T switch in the reference RF front end is to equalize the RF path lengths among all antenna/RF receiver channels.

RADOME

HORN

D

THIS IS THE
D CHOSEN FOR
THE C-BAND ARRAY

WITHOUT
RADOME

HORN-TO-HORN
ISOLATION (dB)

30

40

50

4   3   2   1   0

HORN TO RADOME SPACING D (in.)
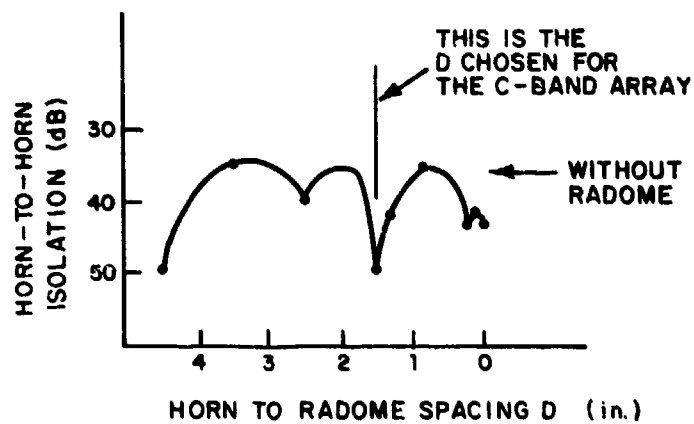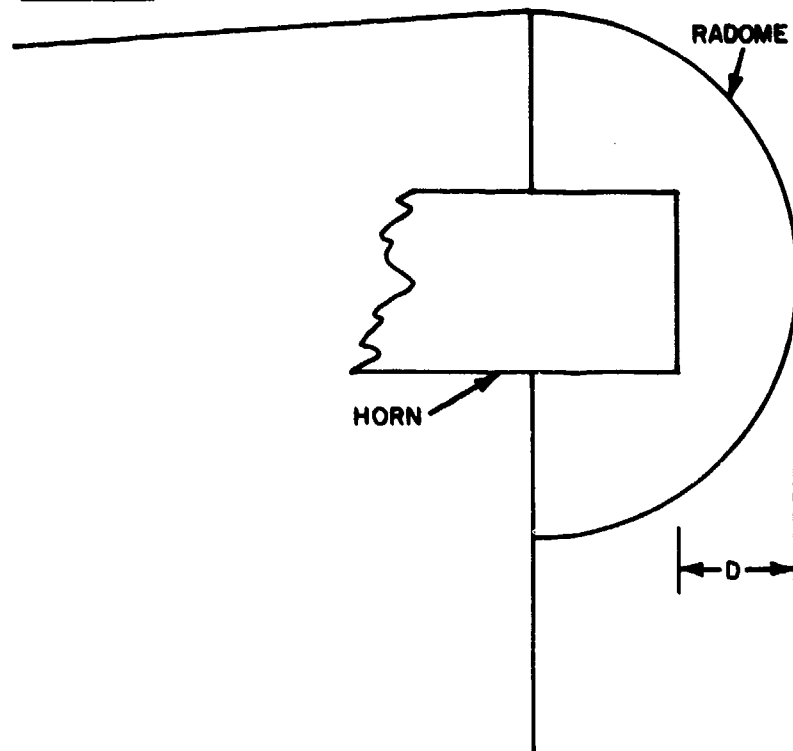
Fig. A-3.   Horn-to-horn isolation versus distance to radome.

109548-N

TYPE N CONNECTOR

|← 1.5 in. →|

3.4 in OD

2 × 1 in. GUIDE

SIDE VIEW

2 in.

3.4 in.
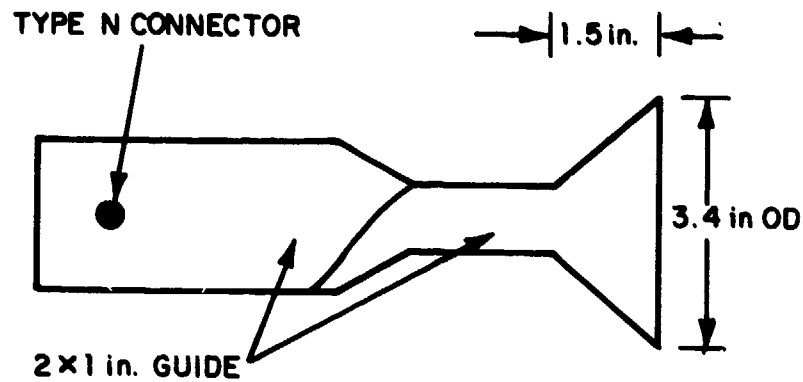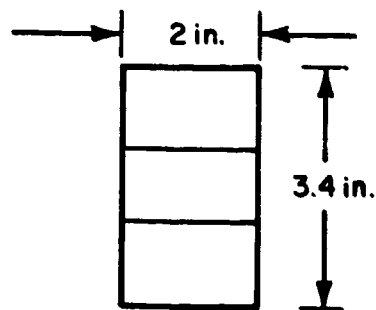
FRONT VIEW

Fig. A-4. Design dimensions for the C-band E-plane sectoral horn.

A-6

E-plane 3-dB beamwidth ≈ 32°

H-plane 3-dB beamwidth ≈ 72°

directivity ≈ 12 dB

maximum E-plane sidelobe level ≈ -12 dB

maximum H-plane sidelobe level ≈ -26 dB

The VSWR with respect to the 50-Ω line for the individual horn assembly (i.e., an E-plane sectoral horn and a twisted waveguide section) was specified to be less than 1.2:1.

EXTERNAL AND RF CALIBRATION DATA

This appendix presents the external and the associated RF calibration data obtained from the array collimation measurements at the Lincoln Laboratory antenna test range during the past year (December 1979 to August 1980). The setup for the array collimation measurements is sketched in Fig. B-1, with a photo of the C-band antenna system positioned for collimation shown in Fig. B-2.[*] The position of the far-field calibration source was fixed. The alignment of the array boresight direction to the far-field source was obtained at the beginning of each measurement day by adjusting the array position such that the array phase center and the array axis fell on the presurveyed alignment point/line marks on the ground at the receiving end. Some of the components (such as RF cables, RF front ends, etc.) in our antenna systems, especially in the L-band antenna system, were repaired or replaced during the past year. Therefore, especially in the L-band calibration data, we can expect to see large changes in the array collimation/RF calibration values from one measurement mission to the next when the components were repaired or replaced in between two measurement missions.

Tables B-1 and B-2 show the calibration data for the C-band antenna system, and Tables B-3 to B-5 are for the L-band antenna system. In each table, two sets of calibration values, taken at two different times (usually they are about 4 to 6 hours apart) of the same mission day, are given to show the short-term variation. In general, our antenna systems remain fairly stable, except for the L-band azimuth array which we have discussed in Section 6-2. The slight variations seen in the C-band array collimation values which cannot be explained by the corresponding RF calibration data are believed to be due to the slight change in the C-band array boresight alignment with the far-field calibration source, either due to the wind blowing

---

[*]The space behind the steel door in the photo is an outdoor anechoic chamber. During the measurement, the steel door was kept open to have the absorption walls of the anechoic chamber behind our C-band antenna system.

(a) C-BAND ANTENNA SYSTEM

(b) L-BAND ANTENNA SYSTEM
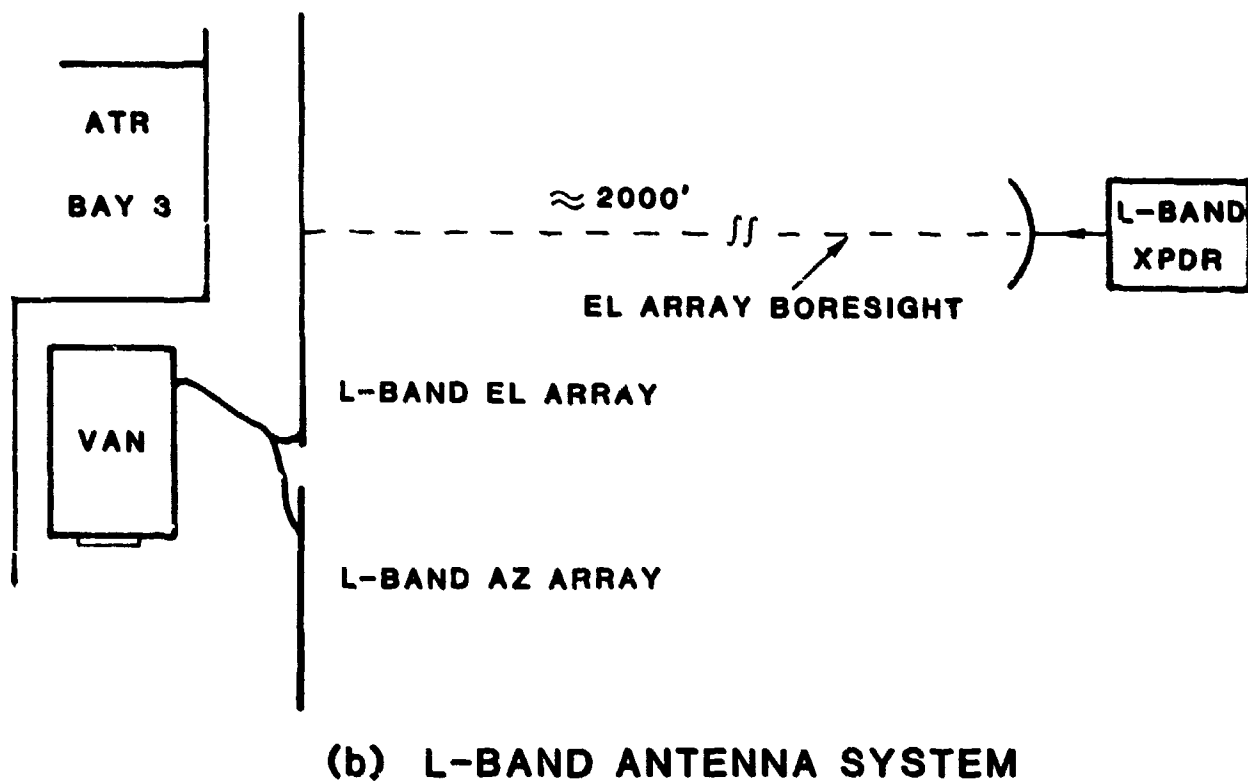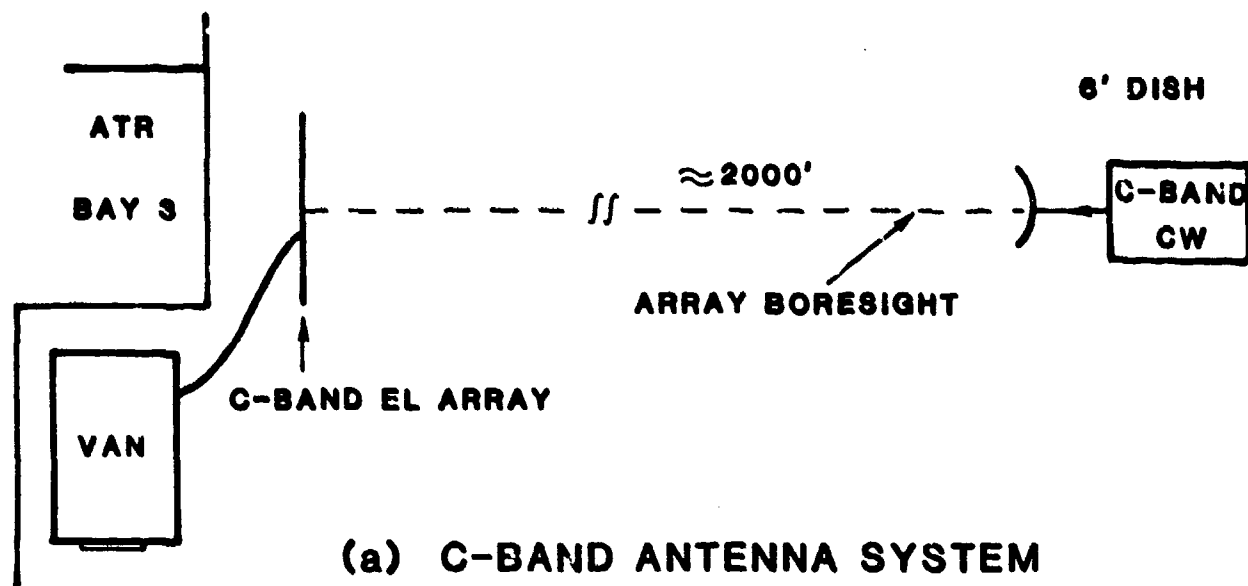
Fig. B-1.   Setup for array collimation measurement.

Fig. B-2. Array collimation measurement for the C-band antenna system.

TABLE B-1

EXTERNAL AND RF CALIBRATION VALUES:
C-BAND ANTENNA SYSTEM, 12/14/79

(a) External calibration values
AMP^1/AMPR and $PH1$/PHR: relative amplitude and phase values from two
external calibrations taken 3 hours apart

ANT.	AMP1	AMPR	AMP1-AMPR	PH1	PHR	PH1-PHR
2	-1.50	-1.25	-0.25	239.92	236.11	3.81
3	-2.25	-2.00	-0.25	209.48	204.73	3.75
4	-3.60	-2.60	-1.00	239.71	235.91	3.80
5	-2.85	-2.75	-0.10	222.70	220.30	2.40
6	-2.25	-2.00	-0.25	226.08	223.22	3.46
7	-2.25	-2.00	-0.25	243.25	244.90	3.27
8	-2.70	-2.30	-0.40	258.22	254.91	3.37
9	-2.25	-2.00	-0.25	267.61	265.17	2.44
10	-2.77	-2.60	-0.17	287.65	284.89	2.76
11	-2.25	-1.90	-0.35	277.44	275.67	1.77
12	-2.25	-2.00	-0.25	204.38	203.35	1.03
13	-0.75	-0.50	-0.25	296.60	295.54	1.06
14	-2.25	-2.00	-0.25	277.36	276.15	1.21
15	-3.60	-2.75	-0.25	255.75	254.64	1.11
16	-0.75	-0.50	-0.25	261.41	259.72	1.69
17	-0.75	-0.50	-0.25	293.37	291.25	2.12
18	-1.50	-1.25	-0.25	290.29	288.46	1.83
19	-1.50	-1.25	-0.25	303.16	303.10	0.06
20	-1.05	-0.95	-0.10	294.10	293.46	0.64
21	-0.98	-0.95	-0.03	320.02	318.46	1.56
22	-0.75	-0.50	-0.25	300.46	299.17	1.29
23	0.75	1.00	-0.25	259.25	258.44	0.81
24	1.50	1.00	0.50	268.10	267.28	0.82
25	-0.68	-0.50	-0.17	249.02	249.06	-0.04
26	-0.75	-0.50	-0.25	256.05	256.14	-0.09
27	0.75	1.00	-0.25	277.13	277.01	0.12
28	0.37	0.62	-0.25	283.17	282.82	0.35
29	-0.75	-0.50	-0.25	229.69	230.24	-0.55
30	0.00	0.00	0.00	0.00	0.00	0.00

(b) RF calibration values
AMP^1/AMPR and PH^1/PHR: relative amplitude and phase values from two RF
calibrations taken 3 hours apart.

ANT.	AMP1	AMPR	AMP1-AMPR	PH1	PHR	PH1-PHR
2	-3.98	-3.24	-0.74	291.05	291.63	-0.58
3	-4.63	-4.40	-0.23	75.09	75.84	-0.74
4	-5.48	-5.40	-0.08	278.29	278.81	-0.52
5	-4.98	-4.75	-0.23	77.17	77.18	-0.00
6	-4.18	-3.90	-0.28	253.63	253.48	0.15
7	-5.27	-4.90	-0.38	108.30	109.58	-1.28
8	-5.18	-4.95	-0.23	279.33	279.26	0.07
9	-3.98	-3.90	-0.08	100.43	100.88	-0.45
10	-3.98	-3.90	-0.08	298.05	297.67	0.38
11	-4.65	-3.98	-0.67	13.14	13.42	-1.28
12	-3.98	-3.90	-0.08	110.86	109.98	0.88
13	-3.98	-3.90	-0.08	216.17	215.36	0.81
14	-5.08	-4.88	-0.18	285.49	285.58	-0.09
15	-5.48	-5.40	-0.08	347.88	347.93	-0.05
16	-3.98	-3.90	-0.08	84.63	84.71	-0.08
17	-3.98	-3.90	-0.08	208.67	209.20	-0.53
18	-3.98	-3.90	-0.08	289.47	289.13	0.34
19	-6.23	-6.15	-0.08	85.22	85.44	-0.22
20	-3.98	-3.90	-0.08	112.84	112.48	0.36
21	-3.98	-3.90	-0.08	282.53	282.60	-0.07
22	-3.56	-3.45	-0.12	288.46	288.26	0.20
23	-0.83	-0.75	-0.07	339.95	341.81	-1.86
24	-2.33	-1.50	-0.82	82.65	83.27	-0.63
25	-4.27	-3.90	-0.38	131.62	133.64	-2.02
26	-3.98	-3.73	-0.24	249.49	249.77	-0.28
27	-2.33	-0.75	-1.58	345.97	347.05	-1.08
28	-2.33	-2.25	-0.07	93.00	95.41	-2.41
29	-3.98	-3.90	-0.08	127.53	127.68	-0.15
30	0.00	0.00	0.00	0.00	0.00	0.00

TABLE B-2

EXTERNAL AND RF CALIBRATION VALUES:
C-BAND ANTENNA SYSTEM, 1/10/80

(a) External calibration values
AMP1/AMPR and PH1/PHR: relative amplitude and phase values from two external calibrations taken 5 hours apart.

ANT.	AMP1	AMPR	AMP1-AMPR	PH1	PHR	PH1-PHR
2	0.33	0.50	-0.18	232.20	231.67	0.53
3	-0.32	0.50	-0.83	198.98	196.12	2.86
4	-0.78	-0.23	-0.55	231.98	230.52	1.46
5	-0.52	0.00	-0.52	216.56	213.87	2.69
6	-0.22	0.50	-0.73	221.26	219.10	2.16
7	-0.37	0.00	-0.37	263.17	263.20	-0.03
8	-0.52	0.00	-0.52	253.37	252.94	0.43
9	0.08	0.60	-0.52	263.63	261.17	2.46
10	-0.83	0.00	-0.83	283.88	281.49	2.39
11	-0.42	0.00	-0.42	274.89	272.94	1.95
12	-0.32	0.40	-0.72	282.51	280.33	2.18
13	0.58	1.05	-0.48	294.60	292.06	2.54
14	-0.83	-0.23	-0.60	275.37	273.02	2.35
15	-1.73	-1.35	-0.38	253.03	251.28	1.75
16	0.63	1.20	-0.57	257.93	254.69	3.24
17	0.68	1.15	-0.47	291.35	289.40	1.95
18	0.08	0.55	-0.47	286.86	285.46	1.40
19	-0.02	0.50	-0.53	301.82	300.54	1.28
20	0.18	0.80	-0.62	290.55	289.53	1.02
21	0.18	0.60	-0.42	317.25	316.27	0.98
22	0.52	0.65	-0.13	299.71	299.41	0.30
23	2.18	2.81	-0.63	262.12	261.76	0.36
24	2.18	3.01	-0.84	271.34	271.56	-0.22
25	0.88	1.05	-0.18	250.12	249.59	0.53
26	1.02	1.50	-0.48	258.98	258.80	0.18
27	2.13	2.51	-0.38	279.00	278.90	0.10
28	1.68	2.00	-0.33	286.23	284.94	1.29
29	0.58	1.00	-0.43	233.51	232.73	0.78
30	0.00	0.00	0.00	0.00	0.00	0.00

(b) RF calibration values
AMP1/AMPR and PH1/PHR: relative amplitude and phase values from two RF calibrations taken 5 hours apart.

ANT.	AMP1	AMPR	AMP1-AMPR	PH1	PHR	PH1-PHR
2	-1.50	0.00	-1.50	291.71	292.13	-0.42
3	-3.04	-1.50	-1.54	77.44	78.76	-1.32
4	-3.90	-2.66	-1.25	278.98	279.59	-0.61
5	-2.88	-1.50	-1.38	79.49	80.99	-1.49
6	-2.05	-0.15	-1.90	254.29	254.19	0.10
7	-3.23	-2.33	-0.90	111.40	111.58	-0.18
8	-3.15	-2.33	-0.82	278.61	278.70	-0.09
9	-1.50	0.00	-1.50	101.16	99.70	1.46
10	-2.33	-0.30	-2.02	298.67	297.13	1.54
11	-2.66	-0.75	-1.91	13.02	11.48	1.54
12	-2.33	-0.75	-1.58	112.76	111.07	1.69
13	-1.50	0.05	-1.55	215.19	213.44	1.75
14	-3.60	-1.27	-2.33	283.88	281.89	1.99
15	-3.90	-1.67	-2.23	348.69	347.70	0.99
16	-2.24	0.00	-2.24	87.19	84.85	2.34
17	-3.15	-1.27	-1.88	209.66	209.63	0.03
18	-2.90	-1.50	-1.40	292.98	291.01	1.97
19	-4.65	-3.15	-1.50	30.88	28.38	2.50
20	-3.15	-0.98	-2.17	114.01	109.58	4.43
21	-2.33	-0.23	-2.10	223.45	221.73	1.72
22	-1.50	0.00	-1.50	291.86	291.02	0.84
23	0.00	1.50	-1.50	350.65	350.25	0.40
24	0.00	1.55	-1.55	91.09	91.96	-0.87
25	-0.75	1.50	-2.25	154.82	153.29	1.53
26	-0.75	0.00	-0.75	254.84	256.58	-1.74
27	0.00	1.50	-1.50	355.63	356.19	-0.56
28	-0.75	1.50	-2.25	101.50	101.32	0.18
29	-1.50	-0.08	-1.42	136.81	134.76	2.05
30	0.00	0.00	0.00	0.00	0.00	0.00

TABLE B-3

EXTERNAL AND RF CALIBRATION VALUES:
L-BAND ANTENNA SYSTEM, 1/24/80

(a) External calibration values:
AMP1/AMPR and PH1/PHR: relative amplitude and phase values
from two external calibrations taken 3 hours apart,

	ANT.	AMP1	AMPR	AMP1-AMPR	PH1	PHR	PH1-PHR
EL array	9	-1.13	-1.05	-0.07	343.64	341.99	1.65
	8	-1.50	-1.50	0.00	347.90	347.60	0.30
	7	-1.13	-1.00	-0.13	22.70	21.72	0.98
	6	-0.15	-0.30	0.15	49.20	48.42	0.78
	5	0.00	-0.75	0.75	244.30	241.34	2.96
	4	0.00	-0.75	0.75	200.17	198.96	1.21
	3.5	-1.50	-0.75	-0.75	218.00	216.53	1.47
	3	0.75	1.50	-0.75	357.87	356.58	1.29
	2.5	0.00	1.50	-1.50	162.61	161.35	1.26
	2	0.00	1.45	-1.45	297.91	296.03	1.88
	1	0.00	0.00	0.00	0.00	0.00	0.00
AZ array	6	-11.40	-9.52	-1.88	60.44	67.67	-7.23
	5	-7.65	-7.05	-0.60	145.65	151.11	-5.46
	4	-7.10	-6.85	-0.25	258.33	264.84	-6.51
	3	-12.93	-12.23	-0.70	342.06	345.97	-3.91
	2	-7.95	-7.50	-0.45	66.13	65.91	0.23
	1	-8.10	-7.35	-0.75	18.79	19.30	-0.51

(b) RF calibration values:
AMP1/AMPR and PH1/PHR: relative amplitude and phase values from
two RF calibrations taken 3 hours apart,

	ANT.	AMP1	AMPR	AMP1-AMPR	PH1	PHR	PH1-PHR
EL array	9	-0.05	-0.65	0.60	20.78	20.98	-0.21
	8	-0.05	-0.65	0.60	20.78	20.98	-0.21
	7	-0.05	-0.65	0.60	20.78	20.98	-0.21
	6	-0.05	-0.65	0.60	20.78	20.98	-0.21
	5	0.00	-0.90	0.90	278.94	277.36	1.58
	4	0.00	-0.90	0.90	278.94	277.36	1.58
	3.5	0.00	-0.90	0.90	278.94	277.36	1.58
	3	0.00	-0.90	0.90	278.94	277.36	1.58
	2.5	0.38	-0.68	1.05	229.31	228.35	0.96
	2	0.38	-0.68	1.05	229.31	228.35	0.96
	1	0.00	0.00	0.00	0.00	0.00	0.00
AZ array	6	-4.65	-6.23	1.58	189.46	204.23	-14.77
	5	-4.65	-6.23	1.58	189.46	204.23	-14.77
	4	-4.65	-6.23	1.58	189.46	204.23	-14.77
	3	-4.65	-6.23	1.58	189.46	204.23	-14.77
	2	-7.88	-8.19	0.31	47.97	54.33	-6.36
	1	-7.88	-8.19	0.31	47.97	54.33	-6.36

EXTERNAL AND RF CALIBRATION VALUES:
L-BAND ANTENNA SYSTEM, 7/17/80

(a) External calibration values:
AMP1/AM-R and PH1/PHR: relative amplitude and phase values from
two external calibrations taken 2 hours apart.

	ANT.	AMP1	AMPR	AMP1-AMPR	PH1	PHR	PH1-PHR
EL array	9	-3.41	-3.45	0.05	300.67	299.93	0.74
	8	-2.58	-2.70	0.12	305.54	305.19	0.35
	7	-3.56	-3.45	-0.11	337.31	336.47	0.84
	6	-3.78	-3.75	-0.03	356.97	356.30	0.67
	5	-0.63	-0.59	-0.04	242.19	240.38	1.81
	4	-0.09	0.03	-0.12	196.17	194.45	1.72
	3.5	-0.96	-0.66	-0.30	214.33	212.62	1.71
	3	-0.96	-0.66	-0.30	348.77	347.00	1.77
	2.5	-1.21	-1.44	0.24	146.78	146.12	0.66
	2	-0.45	-0.49	0.04	284.24	283.71	0.53
	1	0.00	0.00	0.00	0.00	0.00	0.00
AZ array	6	-13.71	-13.52	-0.20	282.58	276.89	5.69
	5	-13.43	-13.26	-0.17	329.38	322.79	6.59
	4	-12.55	-12.51	-0.04	74.94	67.86	7.08
	3	-17.03	-17.04	0.00	162.22	155.45	6.77
	2	-10.58	-10.66	0.08	20.18	19.05	1.13
	1	-12.81	-12.47	-0.34	287.72	286.55	1.17

(b) RF calibration values:
AMP1/AMPR and PH1/PHR: relative amplitude and phase values from
two RF calibrations taken 2 hours apart.

	ANT.	AMP1	AMPR	AMP1-AMPR	PH1	PHR	PH1-PHR
EL array	9	-3.30	-3.33	0.03	316.63	312.12	4.51
	8	-3.30	-3.33	0.03	316.63	312.12	4.51
	7	-3.30	-3.33	0.03	316.63	312.12	4.51
	6	-3.30	-3.33	0.03	316.63	312.12	4.51
	5	-0.75	-1.13	0.38	272.87	267.56	5.31
	4	-0.75	-1.13	0.38	272.87	267.56	5.31
	3.5	-0.75	-1.13	0.33	272.87	267.56	5.31
	3	-0.75	-1.13	0.38	272.87	267.56	5.31
	2.5	-1.88	-1.65	-0.23	213.01	209.81	3.20
	2	-1.88	-1.65	-0.23	213.01	209.81	3.20
	1	0.00	0.00	0.00	0.00	0.00	0.00
AZ array	6	-10.13	-10.31	0.19	197.05	201.95	-4.90
	5	-10.13	-10.31	0.19	197.05	201.95	-4.90
	4	-10.13	-10.31	0.19	197.05	201.95	-4.90
	3	-10.13	-10.31	0.19	197.05	201.95	-4.90
	2	-10.78	-10.70	-0.08	165.80	190.32	-24.52
	1	-10.78	-10.70	-0.08	165.80	190.32	-24.52

TABLE B-5

EXTERNAL AND RF CALIBRATION VALUES:
L-BAND ANTENNA SYSTEM, 7/17/80

(a) External calibration values:
AMP1/AMPR and PH1/PHR: relative amplitude and phase values from
two external calibrations taken 4 hours apart.

	ANT.	AMP1	AMPR	AMP1-AMPR	PH1	PHR	PH1-PHR
EL array	9	-3.03	-3.31	0.28	302.64	300.71	1.93
	8	-2.44	-2.56	0.11	308.56	306.42	2.14
	7	-3.52	-3.42	-0.10	342.57	339.11	3.46
	6	-3.63	-3.83	0.20	1.43	357.74	3.69
	5	0.00	-0.40	0.40	243.70	240.61	3.09
	4	0.55	0.20	0.36	197.92	194.71	3.21
	3.5	-0.50	-0.76	0.26	215.76	212.98	2.78
	3	-0.30	-0.66	0.36	351.69	346.94	4.75
	2.5	-1.23	-1.36	0.13	148.43	145.72	2.71
	2	-0.46	-0.43	-0.03	285.88	283.30	2.58
	1	0.00	0.00	0.00	0.00	0.00	0.00
AZ array	6	-12.96	-13.16	0.20	293.95	282.03	11.92
	5	-13.94	-12.86	-1.07	346.02	328.45	17.57
	4	-12.01	-12.03	0.02	89.04	74.74	14.30
	3	-16.99	-16.86	-0.13	173.90	162.26	11.64
	2	-10.94	-10.65	-0.28	15.78	18.70	-2.92
	1	-12.05	-12.42	0.37	279.99	282.87	-2.88

(b) RF calibration values:
AMP1/AMPR and PH1/PHR: relative amplitude and phase values from two
RF calibrations taken 4 hours apart.

	ANT.	AMP1	AMPR	AMP1-AMPR	PH1	PHR	PH1-PHR
EL array	9	-2.48	-2.55	0.08	314.65	310.19	4.46
	8	-2.48	-2.55	0.08	314.65	310.19	4.46
	7	-2.48	-2.55	0.08	314.65	310.19	4.46
	6	-2.48	-2.55	0.08	314.65	310.19	4.46
	5	-0.23	-0.15	-0.08	272.41	268.77	3.64
	4	-0.23	-0.15	-0.08	272.41	268.77	3.64
	3.5	-0.23	-0.15	-0.08	272.41	268.77	3.64
	3	-0.23	-0.15	-0.08	272.41	268.77	3.64
	2.5	-2.25	-2.25	0.00	213.41	210.00	3.41
	2	-2.25	-2.25	0.00	213.41	210.00	3.41
	1	0.00	0.00	0.00	0.00	0.00	0.00
AZ array	6	-10.53	-10.26	-0.27	303.95	327.62	-23.67
	5	-10.53	-10.26	-0.27	303.95	327.62	-23.67
	4	-10.53	-10.26	-0.27	303.95	327.62	-23.67
	3	-10.53	-10.26	-0.27	303.95	327.62	-23.67
	2	-10.38	-10.06	-0.32	274.84	321.48	-46.64
	1	-10.38	-10.06	-0.32	274.84	321.48	-46.64

during a given measurement day or misalignment of the array boresight to start with on a given measurement day. The case of the large variations sometimes observed between two sets of the C-band array RF calibration data have been conjectured in Section 6-2.

APPENDIX C

SIMULATION RESULTS ON THE EFFECTS OF MEASUREMENT PHASE
AND AMPLITUDE ERRORS ON THE AR AND ARM ALGORITHMS

This appendix presents some simulation results to show the effect of the measurement phase and amplitude errors on the resolution capability of the AR and ARM techniques. The phase (or amplitude) errors for all sensor channels are modeled as the independent random variables with mean errors equal to the desired measured errors. The received signal consists of two plane-wave signals. Signal #1 arrives at -0.5° with respect to the receiving array boresight and signal #2 is at 0°. This corresponds to 0.5-beamwidth separation of two signals for the size of the receiving array used to process the data. Signal #2 is at 3 dB lower power than signal #1, i.e., $P_1/P_2 = 3$ dB. The array signal-to-noise ratio (SNR) is 43 dB. The relative RF phase of two signals is ϕ.

Figure C-1 shows the results for the case of no measurement errors, i.e., phase error (FPHASE) = 0 and amplitude error (FAMP) = 0. Figure C-2 shows the results for the case of phase error only with FPHASE = 3°. Figure C-3 shows the results for the case of amplitude error only with FAMP = 1 dB. For each case, results for three relative phases (ϕ = 0°, 90°, 180°) are given. For each ϕ, 20 spectral estimates are obtained for 20 random noise and error sample functions used. We can see that the 1-dB amplitude error has more undesirable effects on the AR and ARM estimates, in terms of signal resolution and signal arriving angle estimation, than the 3° phase error.

Figure C-4 shows the results for the case of much larger phase error (FPHASE = 5°), and Fig. C-5 shows the results for the case of much larger amplitude error (FAMP = 3 dB). It appears that the 3-dB amplitude error is too severe for the AR and ARM techniques to perform properly. For 0° relative phase case, they cannot even resolve two signals. However, the results for the 5° phase error are still fairly comparable to those for the 3° phase error.

2 plane wave arrivals

Separation angle = 0.5 BW
Power ratio = 3 dB
Array SNR = 43 dB

Relative RF phase between
two signals = φ

(a) φ = 0°

(b) φ = 90°

(c) φ = 180°
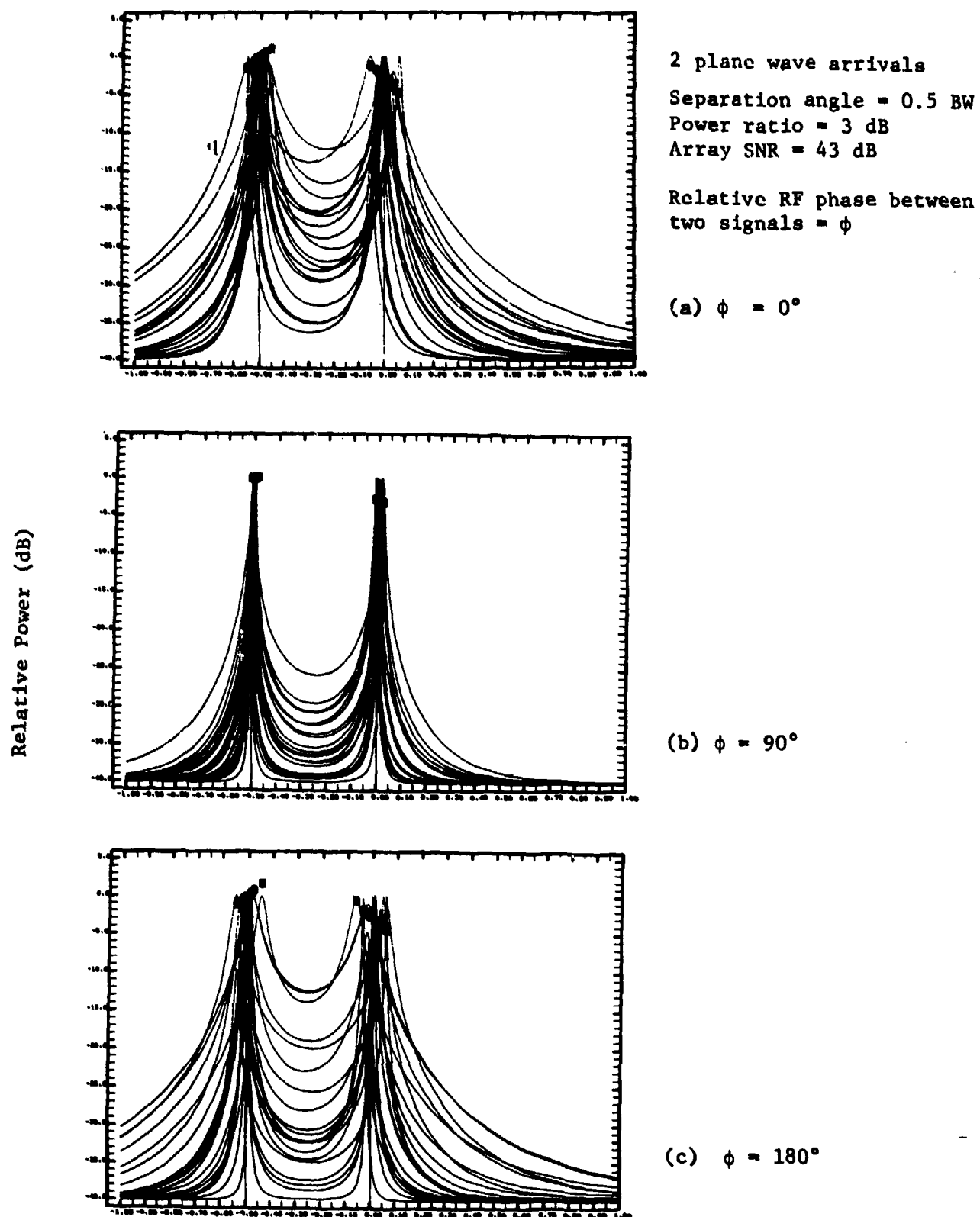
Fig. C-1. AR angular power spectra: Simulation results for
no measurement phase and amplitude errors.
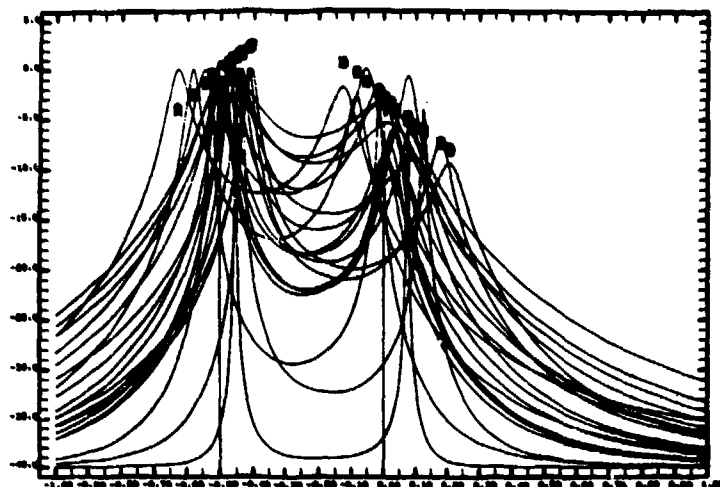
2 plane wave arrivals

Separation angle = 0.5 BW
Power ratio = 3 dB
Array SNR = 43 dB

Relative RF phase between
two signals = φ

(a) φ = 0°

(b) φ = 90°

(c) φ = 180°

Fig. C-2.   AR   angular power spectra: Simulation results
for measurement phase error of 3° without amplitude error.

Fig. C-3. AR angular power spectra: Simulation results for
measurement amplitude error of 1 dB without phase error.

2 plane wave arrivals

Separation angle = 0.5 BW
Power ratio = 3 dB
Array SNR = 43 dB

Relative RF phase between
two signals = ϕ

(a) $\phi = 0°$

(b) $\phi = 90°$

(c) $\phi = 180°$

Fig. C-4. AR angular power spectra: Simulation results
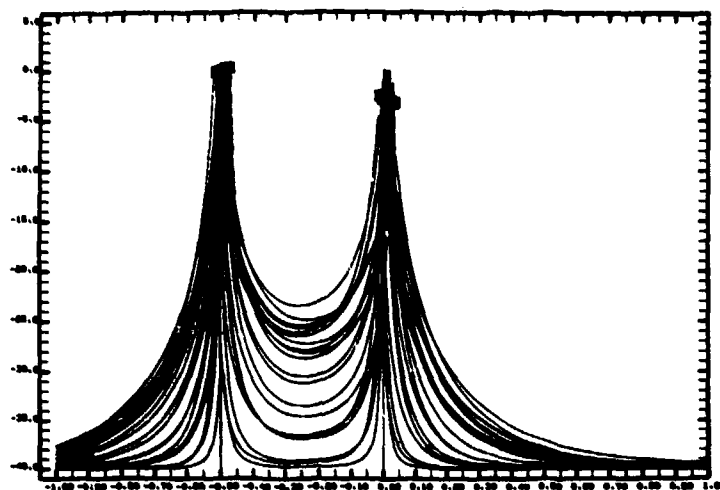for measurement phase error of 5° without amplitude
error.

Relative power (dB)
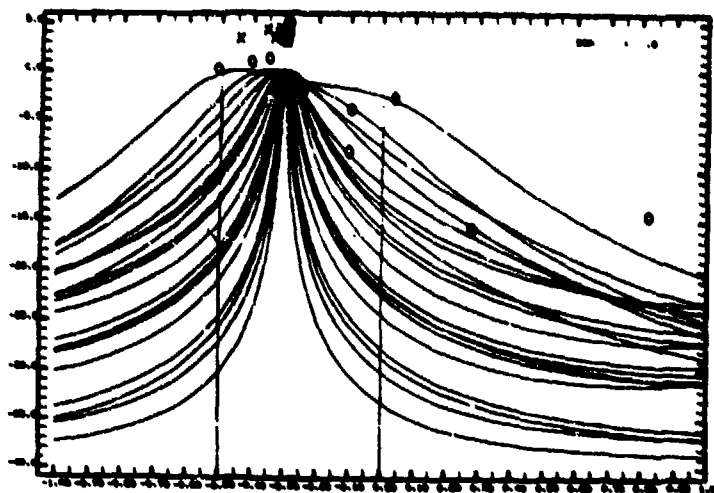
2 plane wave arrivals

Separation angle = 0.5 BU
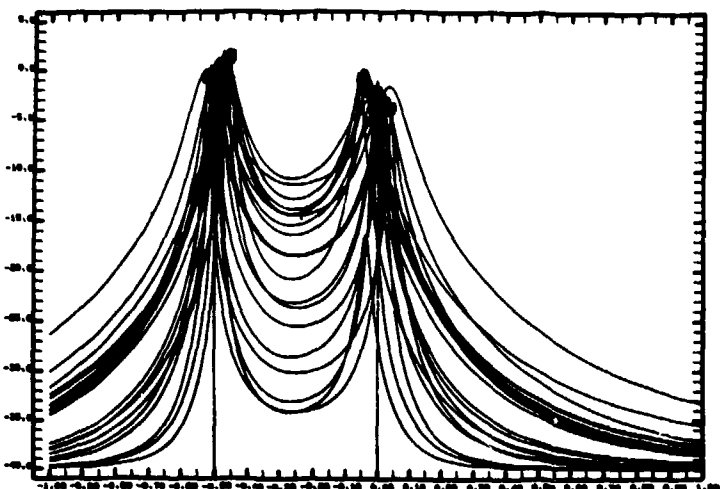Power ratio = 3 dB
Array SNR = 43

Relative RF phase between
two signals = φ

(a) φ = 0°

(b) φ = 90°
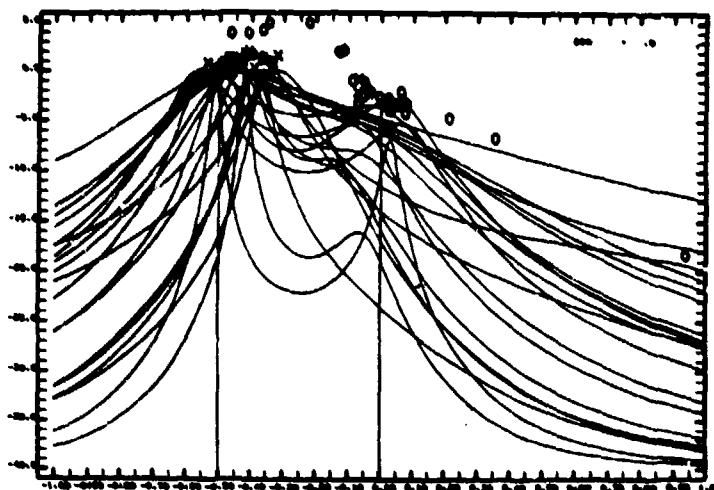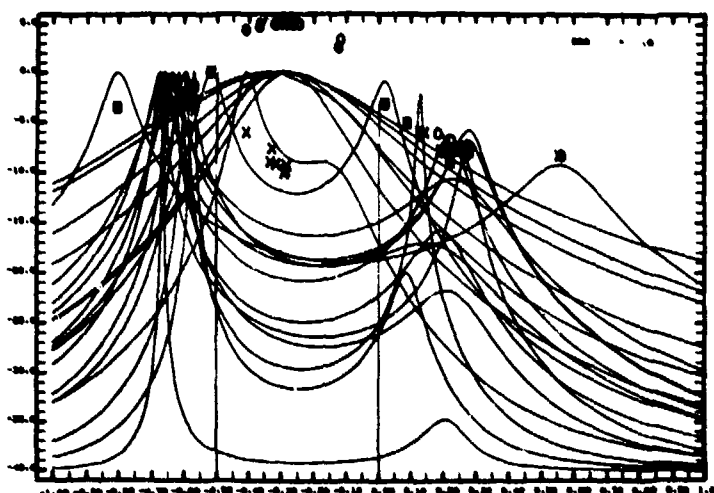
(c) φ = 180°

Fig. C-5. AR angular power spectra: Simulation results
for measurement amplitude error of 3 dB without
phase error.

Table C-1 summarizes results shown in the above figures. It shows some statistics on two-signal resolution with various measurement phase and amplitude errors.

TABLE C-1
STATISTICS FOR TWO SIGNAL RESOLUTION WITH
VARIOUS MEASUREMENT PHASE AND AMPLITUDE ERRORS

CASE			φ = 0°		φ = 90°		φ = 180°	
			AR	ARM	AR	ARM	AR	ARM
		% DET	100	100	100	100	100	100
FPHASE = 0°	#1	x̄	-0.003	-0.004	-.0002	-.0004	-0.002	-0.003
FAMP = 0 DB		S	0.023	0.023	0.006	0.006	0.022	0.023
(FIG. C-1)	#2	x̄	-0.002	-0.002	0.002	0.003	0.004	0.004
		S	0.029	0.029	0.009	0.009	0.03	0.03
		% DET	100	100	100	100	100	100
FPHASE = 3°	#1	x̄	0.015	0.014	0.004	0.004	0.01	0.009
FAMP = 0 DB		S	0.052	0.052	0.009	0.01	0.022	0.022
(FIG. C-2)	#2	x̄	0.031	0.035	0.009	0.009	-0.013	-0.012
		S	0.079	0.079	0.014	0.013	0.03	0.03
		% DET	35	100	95	100	90	100
FPHASE = 0°	#1	x̄	0.017	0.029	-0.002	0.001	-0.014	0.004
FAMP = 1 DB		S	0.023	0.152	0.028	0.034	0.071	0.09
FIG. (C-3)	#2	x̄	-0.012	-0.019	0.001	0.0004	0.024	0.002
		S	0.031	0.243	0.032	0.04	0.097	0.12
		% DET	70	100	100	100	100	100
FPHASE = 5°	#1	x̄	0.033	0.02	0.006	0.006	0.012	0.009
FAMP = 0°		S	0.072	0.097	0.013	0.013	0.031	0.03
(FIG. C-4)	#2	x̄	0.071	0.049	0.012	0.014	-0.015	-0.012
		S	0.136	0.136	0.018	0.018	0.044	0.042
		% DET	0	40	25	100	65	100
FPHASE = 0°	#1	x̄	--	0.15	-0.006	0.037	-0.139	-0.035
FAMP = 3 DB		S	--	0.069	0.05	0.06	0.065	0.154
(FIG. C-5)	#2	x̄	--	0.138	-0.003	0.051	0.216	0.065
		S	--	0.398	0.057	0.237	0.122	0.241

% DET: Percentage of successful resolution (i.e., existence of 2 spectral
peaks > -15 dB)
#1: Signal #1,   #2: Signal #2
x̄: Mean angular error in degrees in locating signal #1, i = 1,2
S: Standard deviation of angular error in degree.

APPENDIX D

MAXIMUM ENTROPY SPECTRAL ESTIMATION, AUTOREGRESSIVE PROCESSES, AND LINEAR PREDICTION

In this appendix, we investigate the principal relationships between the information theoretic concepts of mutual information and entropy, autoregressive processes, and linear prediction. In the course of the discussion, we will develop the key theoretical ideas behind an area of research that has come to be widely known as maximum entropy spectral analysis. While none of the material presented here is entirely original, perhaps some modest benefit will accrue by exploring, in a logical and mathematically consistent manner, the important relationships that exist between the underlying concepts mentioned above.

Information and Entropy

The average mutual information between two (continuous) random vectors $\underline{r}$ and $\underline{s}$ may be defined [90] to be

$$\overline{\imath}(\underline{r};\underline{s}) \overset{\Delta}{=} E\left\{\ln \frac{\psi_{\underline{r},\underline{s}}}{\psi_{\underline{r}} \cdot \psi_{\underline{s}}}\right\}$$

$$= \overline{\imath}(\underline{s};\underline{r}) \tag{D.01}$$

where ψ generally denotes a probability density function (pdf).

The statistical average (expected value) is taken with respect to the joint pdf $\psi_{\underline{r},\underline{s}}$ of $\underline{r}$ and $\underline{s}$. For our purposes here, it is convenient to use the natural algorithm, and information is measured in nats (or nits) rather than bits. The average information is zero if $\underline{r}$ and $\underline{s}$ are statistically independent. In general, $\underline{r}$ and $\underline{s}$ are statistically dependent, and the conditional pdf

$$\psi_{\underline{r}|\underline{s}} \overset{\Delta}{=} \psi_{\underline{s},\underline{r}} / \psi_{\underline{r}} \tag{D.02}$$

is useful, i.e.,

$$\overline{\imath}(\underline{s};\underline{r}) = E\{\ln \psi_{\underline{r}|\underline{s}}\} - E\{\ln \psi_{\underline{r}}\} . \tag{D.03}$$

When the (observed) vector $\underline{r}$ is the sum of the (signal) vector $\underline{s}$ and a statistically indpendent (noise) vector $\underline{n}$, the conditional pdf may be evaluated in terms of the pdf of $\underline{n}$, i.e.,

$$\psi_{\underline{r}|\underline{s}} (\underline{y}|\underline{x}) = \psi_{\underline{n}} (\underline{x} - \underline{y}) . \tag{D.04}$$

It follows easily that

$$E\{\ln \psi_{\underline{r}|\underline{s}}\} = \iint [\ln \psi_{\underline{r}|\underline{s}}(\underline{y}|\underline{x})] \psi_{\underline{r},\underline{s}}(\underline{x},\underline{y}) \, d\underline{x} \, d\underline{y}$$

$$= \iint [\ln \psi_{\underline{n}}(\underline{x} - \underline{y})] \psi_{\underline{n}}(\underline{x} - \underline{y}) \psi_{\underline{s}}(\underline{y}) \, d\underline{x} \, d\underline{y} .$$

The obvious change of variables $\underline{x}' = \underline{x} - \underline{y}$, $\underline{y}' = \underline{y}$ has a Wronskian determinant of unity. Consequently,

$$E\{\ln \psi_{\underline{r}|\underline{s}}\} = \iint [\ln \psi_{\underline{n}}(\underline{x}')] \psi_{\underline{n}}(\underline{x}') \psi_{\underline{s}}(\underline{y}') \, d\underline{x}' \, d\underline{y}'$$

$$= \int [\ln \psi_{\underline{n}}(\underline{x}')] \psi_{\underline{n}}(\underline{x}') \, d\underline{x}' \int \psi_{\underline{s}}(\underline{y}') \, d\underline{y}'$$

$$= E\{\ln \psi_{\underline{n}}\} . \tag{D.05}$$

Thus, the average information about a signal $\underline{s}$ contained in a noisy observation $\underline{s} + \underline{n}$ is

$$\imath(\underline{s}; \underline{s} + \underline{n}) = E\{\ln \psi_{\underline{n}}\} - E\{\ln \psi_{\underline{r}}\} . \tag{D.06}$$

This result may be expressed in terms of the entropy functional

$$\eta(\psi) \triangleq - \int [\ln \psi(\underline{x})] \psi(\underline{x}) \, d\underline{x} \tag{D.07}$$

D-2

where ψ is any (marginal) pdf and the integral is over the domain of ψ, i.e.,

$$\iota(\underline{s};\ \underline{s} + \underline{n}) = \eta(\psi_{\underline{s} + \underline{n}}) - \eta(\psi_{\underline{n}}) \tag{D.08}$$

for statistically independent vectors $\underline{s}$ and $\underline{n}$.

The result in Eq. (D.08) shows that the average information contained in a noisy observation of a desired signal is the entropy of the received signal less the entropy of the noise. We may conclude that maximizing the received entropy (given such power) yields the maximum average information, provided the signal is received in statistically independent noise (beyond our control).

Gaussian Signals in Gaussian Noise

In principle, the entropy of a random vector can always be obtained from Eq. (D.07), given the pdf of the vector. Fortunately, the pdf for Gaussian vectors is easy to specify, and _Gaussian_ entropy is not difficult to calculate. Here, we consider the case of a (circular) complex Gaussian random vector $\underline{g}$ with mean $\underline{m}$ and (non-singular) covariance $\underline{C}$, i.e.,

$$\underline{m} = E\{\underline{g}\} \tag{D.09}$$

and

$$\underline{C} = E\{(\underline{g} - \underline{m})\ (\underline{g} - \underline{m})^H\} \tag{D.10}$$

where $\underline{C}$ is a positive-definite Hermitian matrix.

The concept of a complex vector is seldom useful unless its real and imaginary random components are identically distributed and uncorrelated, i.e.,

$$E\{(\underline{g} - \underline{m})\ (\underline{g} - \underline{m})^T\} = \underline{0} \qquad \text{(circular } \underline{g}) \ .$$

The pdf of a Gaussian random vector satisfying the above conditions is

$$\psi_{\underline{g}}(\underline{x}) = \frac{1}{(2\pi)^L |\underline{C}|} \; e^{-(\underline{x} - \underline{m})^H \underline{C}^{-1} (\underline{x} - \underline{m})}. \qquad (D.11)$$

where L is the dimension of $\underline{g}$ and $|\underline{C}|$ denotes the determinant of $\underline{C}$. The Gaussian entropy is

$$\eta(\psi_{\underline{g}}) = -E\{\ell n \; \psi_{\underline{g}}\}$$
$$= E\{\ell n (2\pi)^L |\underline{C}|\} + E\{(\underline{x} - \underline{m})^H \underline{C}^{-1} (\underline{x} - \underline{m})\} \; .$$

The expression inside the first pair of brackets is a constant, and the second integrates to L, e.g.,

$$E\{(\underline{x} - \underline{m})^H \underline{C}^{-1} (\underline{x} - \underline{m})\} = E\{tr[\underline{C}^{-1} (\underline{x} - \underline{m}) (\underline{x} - \underline{m})^H]$$
$$= tr\{\underline{C}^{-1} E[(\underline{x} - \underline{m}) (\underline{x} - \underline{m})^H]\}$$
$$= tr\{\underline{C}^{-1} \underline{C}\}$$
$$= L \; .$$

Consequently, the entropy of an L-dimensional complex Gaussian vector does not depend on its mean $\underline{m}$, i.e.,

$$\eta(\psi_{\underline{g}}) = L(1 + \ell n \; 2\pi) + \ell n |\underline{C}| \; . \qquad (D.12)$$

When a Gaussian (signal) vector is observed in statistically independent (s.i.) Gaussian noise, the average information per element (sample) is given by using Eq. (D.12) in Eq. (D.08) and dividing the result by L. Thus, we let

$$\gamma_L \overset{\Delta}{=} \frac{1}{L} (\ell n |\underline{S} + \underline{N}| - \ell n |\underline{N}|) \qquad (D.13)$$
$$= \frac{1}{L} \; \ell n |\underline{N}^{-1} \underline{S} + \underline{I}|$$

denote the information rate for a Gaussian signal with (non-singular) covariance $\underline{S}$ observed in Gaussian noise with (non-singular) covariance $\underline{N}$. In the following sections, we will deal only with the normalized white noise case,

D-4

i.e., $\underline{N} = \underline{I}$. In this important special case, the Gaussian information rate is

$$
\begin{aligned}
\gamma_L &= \frac{1}{L} \ell n |\underline{S} - \underline{I}| \\
&= \frac{1}{L} \ell n \prod_{\ell=1}^{L} (1 + \lambda_\ell) \\
&= \frac{1}{L} \sum_{\ell=1}^{L} \ell n (1 + \lambda_\ell)
\end{aligned}
\tag{D.14}
$$

where $\{\lambda_\ell | \ell=1,\ldots,L\}$ are the eigenvalues of $\underline{S}$. The general case is handled similarly by finding the eigenvalues of $\underline{S}$ in the metric of $\underline{N}$; i.e., the characteristic equation for the eigenvalues is always

$$
|\underline{S} - \lambda \underline{N}| = 0 \, . \, . \tag{D.15}
$$

As the number of samples increases, the limiting information rate per sample is written as

$$
\gamma_\infty \overset{\Delta}{=} \lim_{L \to \infty} \gamma_L \, . \tag{D.16}
$$

In the next section, the limiting function rate is examined for the important class of stationary Gaussian processes.

Stationary Gaussian Processes

When the signal covariance matrix $\underline{S}$ is Toeplitz, i.e.,

$$
(\underline{S})_{mn} = c_{m-n} \quad , \tag{D.17}
$$

the (zero-mean) signal $\underline{s}$ is said to be (wide-sense) stationary. Since a covariance matrix is always Hermitian, the correlation lags satisfy the (harmonic) relationship

$$
c_{-m} = c_m^* \tag{D.18}
$$

D-5

an. hence the power spectral density

$$\chi(\alpha) \overset{\Delta}{=} \sum_{-\infty}^{\infty} c_m e^{-im\alpha} \quad , \tag{D.19}$$

is a real function. More generally, the z-transform of the lags is written as

$$C(z) \overset{\Delta}{=} \sum_{-\infty}^{\infty} c_m z^{-m} . \tag{D.20}$$

We refer to a positive semi-definite matrix that is both Hermitian and Toeplitz as a _correlation_ matrix.

As the number of lags increases without limit, the eigenvectors of a correlation matrix must approach (orthogonal) eigenfunctions of

$$\sum_{m=-\infty}^{\infty} c_{n-m} e_m = \tilde{\lambda} e_n \tag{D.21}$$

The solutions of Eq. (D.21) take the form

$$e_n = kz^n \tag{D.22}$$

for any complex numbers z and k. The eigenvalue associated with z is given by the z-transform of the lags, i.e.,

$$\tilde{\lambda}(z) = C(z) . \tag{D.23}$$

For z on the unit circle, the (limiting) eigenvalues are given by the signal's (real) power spectral density, i.e.,

$$\tilde{\lambda} (e^{i\alpha}) = \chi(\alpha) . \tag{D.24}$$

Consequently, we consider approximations of the form

$$\underline{\tilde{S}} = \underline{V}(\underline{\alpha}_L) \ \underline{\tilde{\Lambda}}(\underline{\alpha}_L) \ \underline{V}^H(\underline{\alpha}_L) , \tag{D.25}$$

where $\underline{V}(\underline{\alpha}_L)$ is a (square) direction matrix of the type of introduced in

D-6

Chapter II, and $\underline{\tilde{\Lambda}}(\underline{\alpha}_L)$ is a diagonal matrix of eigenvalues obtained by evaluating the power spectral density (D.24) at uniform intervals around the unit circle, i.e., the elements of $\underline{\alpha}_L$ are

$$\alpha_\ell = 2\pi\ell/L \qquad ; \ell = 1,\ldots,L .\tag{D.26}$$

Assuming the approximation in Eq. (D.25) converges to $\underline{S}$ in some suitable sense as $L \to \infty$, the limiting information rate Eq. (D.16) for a stationary Gaussian process in white Gaussian noise is (evidently) given by

$$
\begin{aligned}
\gamma_\infty &= \lim_{L \to \infty} \gamma_L \\[2mm]
&= \lim_{L \to \infty} \frac{1}{L} \ln|\underline{\tilde{S}}| \\[2mm]
&= \lim_{L \to \infty} \frac{1}{L} \ln|\underline{\tilde{\Lambda}}(\underline{\alpha}_L)| \\[2mm]
&= \lim_{L \to \infty} \frac{1}{L} \sum_{\ell=1}^{L} \ln \tilde{\lambda}(e^{i\alpha_\ell}) \\[2mm]
&= \lim_{L \to \infty} \frac{1}{L} \sum_{\ell=1}^{L} \ln \chi(\alpha_\ell) \\[2mm]
&= \frac{1}{2\pi} \int_{-\infty}^{\infty} \ln \chi(\alpha)\, d\alpha .
\end{aligned}
\tag{D.27}
$$

The absolute integrability of the logarithm of the power spectral density is known as the Paley-Weiner condition. Under this condition, Szego [90] derives Eq. (D.27) in a mathematically correct manner. Our _heuristic_ approach is offered simply as a plausibility argument. The interested (brave) reader is referred to the first five chapters of [90] for a rigorous treatment of the subject.

Maximum Entropy

In this secton, we first show that the Gaussian pdf maximizes the entropy functional for random vectors with specified (i.e., known) mean and covariance. A similar type of argument then leads to the 'maximum entropy' power

spectral density for a stationary Gaussian random process. In both cases, the discussion is based on certain rudimentary results from elementary functional analysis. In particular, at least a nodding acquaintance with the basic properties of convex sets and functions is assumed. A good treatment of these topics may be found in Luenberger [91].

Maximum Entropy Probability Models

In a previous section, the entropy of a random vector was <u>defined</u> in terms of its pdf. Quite often, only second-order statistics of a random signal (vector) are available. These statistics may be interpreted as constraints on the set of <u>admissible</u> probability models. Specifically, an admissible model for a random signal $\underline{s}$ with specified mean $\underline{m}$ and non-singular covariance $\underline{C}$ is defined to be any (real) function ψ that is non-negative over the domain of the signal and, in addition, satisfies

$$\int \psi(\underline{s}) \, d\mu(\underline{s}) = \underline{1} \tag{D.28a}$$

$$\int \underline{s} \, \psi(\underline{s}) \, d\mu(\underline{s}) = \underline{m} \tag{D.28b}$$

and

$$\int (\underline{s} - \underline{m})(\underline{s} - \underline{m})^H \psi(\underline{s}) \, d\mu(\underline{s}) = \underline{C} \tag{D.28c}$$

where μ is the standard (i.e., Lebesgue) measure in the signal domain. Of course, the integrals in Eq. (D.28) are evaluated over the entire signal domain (e.g., a finite-dimensional complex vector space).

Now suppose that ψ and $\psi + \delta\psi$ are two distinct admissible models. Since the set of admissible models is clearly convex,

$$(1 - \tau)\psi + \tau(\psi + \delta\psi) = \psi + \tau\delta\psi \quad ; \quad 0 < \tau < 1$$

is also an admissible model, and we say that $\delta\psi$ is an <u>admissible perturbation</u>. From Eq. (D.28), we conclude that any admissible perturbation $\delta\psi$ must satisfy

$$\int \delta\psi(\underline{s}) \, d\mu(\underline{s}) = 0 \tag{D.29}$$

$$\int (\underline{s} - \underline{m}) (\underline{s} - \underline{m})^H \delta\psi(\underline{s}) \, d\mu(\underline{s}) = 0 \qquad (D.30)$$

as well as

$$\int \underline{s} \, \delta\psi(\underline{s}) \, d\mu(\underline{s}) = 0 .$$

Multiplying Eq. (D.30) by $\underline{C}^{-1}$ and taking the trace of the result gives

$$\int (\underline{s} - \underline{m})^H \underline{C}^{-1} (\underline{s} - \underline{m}) \, \delta\psi(\underline{s}) \, d\mu(\underline{s}) = 0 . \qquad (D.31)$$

Of course, Eq. (D.29) remains valid when muliplied by any (non-zero) scalar.

As an immediate consequence, we observe that the (complex) Gaussian density ψ_g has the property

$$\int \delta\psi \, \ell n \, \psi_g \, d\mu = 0 \qquad (D.32)$$

for any admissible perturbation. This important intermediate result follows directly from Eqs. (D.29) and (D.31) by first writing Eq. (D.11) as

$$-\ell n \, \psi_g(\underline{s}) = \ell n (2\pi)^L |\underline{C}| + (\underline{s} - \underline{m})^H \underline{C}^{-1} (\underline{s} - \underline{m}) .$$

Multiplying this equation by $\delta\psi(\underline{s})$ and integrating over the signal domain yields the desired result Eq. (D.32).

Recalling Eq. (D.07), we now differentiate

$$\begin{aligned}
\eta(\psi + \tau\delta\psi) &= -E\{\ell n(\psi + \tau\delta\psi)\} \\
&= -\int (\psi + \tau\delta\psi) \, \ell n(\psi + \tau\delta\psi) \, d\mu
\end{aligned}$$

with respect to τ. It then follows from Eq. (D.29) that

$$\frac{\partial}{\partial\tau} \eta(\psi + \tau\delta\psi) = -\int \delta\psi \, \ell n(\psi + \tau\delta\psi) \, d\mu . \qquad (D.33)$$

Differentiating once more yields a fundamental property of entropy,

$$\frac{\partial^2}{\partial \tau^2} \, \eta(\psi + \tau\delta\psi) = -\int (\delta\psi)^2 (\psi + \tau\delta\psi)^{-1} d\mu \; . \qquad (D.34)$$

$$< 0$$

This important inequality shows that the entropy functional is strictly con-cave over any convex domain. Therfore, any (admissible) stationary point of the entropy functional η also maximizes η over the convex set of (admissible) pdfs.

By definition, any (admissible) $\hat{\psi}_g$ is a stationary point of $\eta(\psi)$ iff

$$\lim_{\tau \to 0+} \; \eta(\hat{\psi} + \tau\delta\psi) = 0 \qquad (D.35)$$

for every admissible perturbation $\delta\psi$. Since the complex Gaussian density in (D.11) is obviously admissible, we may substitute ψ_g for ψ in Eq. (D.33). Formally taking the limit as $\tau \to 0^+$ (i.e., from above) then gives

$$\lim_{\tau \to 0+} \; \frac{\partial \eta}{\partial \tau} \, (\psi_g + \tau\delta\psi) = -\int \delta\psi \, \ln \psi_g \, d\mu \; . \qquad (D.36)$$

Comparing Eq. (D.35) and (D.36), we may conclude that ψ_g is the maximum entro-py density iff $\int \delta\psi \, \ln \psi_g \, d\mu = 0$ for every admissible $\delta\psi$; however, this pro-perty has already been established Eq. (D.32) for the Gaussian density ψ_g. Therefore, the maximum entropy pdf for a random vector with specified mean and covariance always Gaussian. Moreover, the numerical value of the maximum entropy does not depend on the mean Eq. (D.12).

Maximum Entropy Spectral Estimation

Given a finite sequence of correlation lags,

$$c_\ell = \frac{1}{2\pi} \int_{-\pi}^{\pi} \chi(\alpha) \, e^{i\alpha\ell} \qquad ; \qquad \ell = 0, 1, \ldots, L \qquad (D.37)$$

the maximum entropy power spectral density $\hat{\chi}$ is defined [92] to be the power spectral density (psd) that maximizes the Gaussian information rate (i.e.,

'entropy')

$$\gamma_\infty(\chi) = \frac{1}{2\pi} \int_{-\pi}^{\pi} \ln \chi(\alpha) \, d\alpha \tag{D.38}$$

subject to the lag constraints in Eq. (D.37). It is assumed here that the correlation (Toeplitz) matrix constructed from the given lags is positive definite. Of course, the remaining lags may be estimated (calculated) from Eq. (D.37) as soon as the maximum entropy (ME) estimate $\hat{\chi}$ is found. As it turns out, the extended ME lags also satisfy a simple recursive relationship [93].

An admissible power spectral density (psd) is any real non-negative function χ that satisfies Eq. (D.37). Clearly, the domain of admissible power spectral densities is convex. From Eq. (D.19) or (D.37), we conclude that any admissible spectral perturbation may be expressed as

$$\delta\chi(\alpha) = \sum_{|n|>L} e^{-in\alpha} \, \delta c_n \, . \tag{D.39}$$

Of course, the corresponding lag perturbations in Eq. (D.39) must satisfy

$$\delta c_{-n} = \delta c_n^* \tag{D.40}$$

in order for $\delta\chi$ to be an admissible (i.e., real) spectral perturbation.

Following essentially the same procedure as before, we first differentiate $\gamma_\infty(\chi + \tau\delta\chi)$ with respect to τ, i.e.,

$$\frac{\partial}{\partial \tau} \gamma_\infty(\chi + \tau\delta\chi) = \frac{1}{2\pi} \int_{-\pi}^{\pi} \frac{\delta\chi}{\chi + \tau\delta\chi} \, d\alpha \, . \tag{D.41}$$

Differentiating again, i.e.,

$$\frac{\partial^2}{\partial \tau^2} \gamma_\infty(\chi + \tau\delta\chi) = -\frac{1}{2\pi} \int_{-\pi}^{\pi} \left(\frac{\delta\chi}{\chi + \tau\delta\chi} \right)^2 \, d\alpha, \tag{D.42}$$

$$< 0$$

we found that $\gamma_\infty(\chi)$ is (also) a concave functional over any convex domain of

(admissible) power spectral densities. Consequently, the ME psd $\hat{\chi}$ exists and is unique.

Substituting $\hat{\chi}$ for χ in Eq. (D.41) and taking the limit as $\tau \rightarrow 0^+$ yields

$$\lim_{\tau \rightarrow 0^+} \frac{\partial}{\partial \tau} \gamma_\infty (\hat{\chi} + \tau \delta \chi) = \frac{1}{2\pi} \int_{-\pi}^{\pi} \hat{\chi}^{-1} \delta \chi \, d\alpha \qquad (D.43)$$
$$= 0$$

since $\hat{\chi}$ must (also) be a stationary point of γ_∞. Comparing Eqs. (D.41) and (D.43), we conclude that the ME psd satisfies

$$\frac{1}{2\pi} \int_{-\pi}^{\pi} \hat{\chi}^{-1}(\alpha) \left(\sum_{|n|>L} e^{in\alpha} \delta c_n \right) d\alpha = 0 \; . \qquad (D.44)$$

Since $\hat{\chi}$ is real, we immediately deduce the equivalent condition(s)

$$f_n \overset{\Delta}{=} \frac{1}{2\pi} \int_{-\pi}^{\pi} \hat{\chi}^{-1}(\alpha) e^{in\alpha} \, d\alpha = 0 \qquad ; \qquad |n| > L \qquad (D.45)$$

for the (admissible) maximum entropy power spectral density. If (D.45) is violated, we may obviously choose $\delta c_n = \epsilon f_n^*$ for some sufficiently small $\epsilon > 0$ and some $n > L$ such that $f_n \neq 0$. We may honor Eq. (D.40) by setting $\delta c_{-n} = \epsilon f_n$, and the remaining lag perturbations are set to zero. We would then calculate

$$\frac{1}{2\pi} \int_{-\pi}^{\pi} \hat{\chi}^{-1} \delta \chi \, d\alpha = 2\epsilon \, |f_n|^2$$

in violation of Eq. (D.43). Thus, the necessity of Eq. (D.45) is established by contradiction.

The result obtained in Eq. (D.45) shows that the inverse of the maximum entropy "spectrum" (power spectral density) has a finite Fourier series. This immediately suggests that the ME signal is an autoregressive process.

In the next two sections, autoregressive processes are defined and some of their fundamental properties examined. In particular, we will show that there is a unique (admissible) autoregressive process of order L that satis-

fies the correlation constraint in Eq. (D.37). Moreover, the power spectral density of this or any other L^{th} order autoregressive process is shown to be of the form

$$\chi_L(\alpha) = \frac{\omega_L}{|P_L(e^{i\alpha})|^2} \qquad (D.46)$$

where $\omega_L > 0$ and $P_L(z)$ is an Lth order polynomial in z of the form

$$P_L(z) = z^L + \ldots + p_L \ .$$

It follows that $\chi_L^{-1}(\alpha)$ is a 'trig' polynomial of order L in $e^{i\alpha}$, i.e.,

$$\omega_L \chi_L^{-1} = |P_L(e^{i\alpha})|^2 \qquad \qquad . \qquad (D.47)$$

$$= e^{-i\alpha L} + \ldots + q_\ell e^{i\alpha\ell} + \ldots + e^{i\alpha L}$$

Integrating Eq. (D.47) over $(-\pi, \pi)$, we observe[*] that

$$\int_{-\pi}^{\pi} \chi_L^{-1}(\alpha)\, e^{in\alpha}\, d\alpha = 0 \qquad ; \qquad |n| > L \qquad (D.48)$$

for any L^{th} order AR power spectral density. However, only one L^{th} order AR process is admissible. Comparing Eq. (D.45) and Eq. (D.48), we conclude that the unique 'maximum entropy' power spectral density consistent with the known correlation lags in Eq. (D.37) is an autoregressive process. In a subsequent section, we will discuss an efficient algorithm for actually calculating the (unique) autoregressive parameters from a given set of lags.

Autoregressive Processes

The concept of an autoregressive process is closely tied to the idea of linear prediction. The development here is based on the usual orthogonal

---

[*]Recall that $\int_{-\pi}^{\pi} e^{in\theta}\, d\theta \neq 0$

decomposition of a (zero-mean) random process $\{s_k\}$. Thus, we consider first the best linear estimate of s_k based on the entire past history of the process, i.e., there exists a linear operator $\hat{L}$ such that

$$\hat{s}_{k|k-1} \overset{\Delta}{=} \hat{L}(s_{k-1}, s_{k-2}, \ldots) \tag{D.49}$$

minimizes the mean square error $E|s_k - \hat{s}_{k|k-1}|^2$ over all (linear) estimates of the general form $L(s_{k-1}, s_{k-2}, \ldots)$.

The 'innovations' process is defined as

$$w_k \overset{\Delta}{=} s_k - \hat{s}_{k|k-1} \cdot \tag{D.50}$$

The well-known _orthogonality principle_ [94]

$$E\{(s_k - \hat{s}_{k|k-1}) s_j^*\} = 0 \quad ; \quad j < k \tag{D.51}$$

guarantees that w_k and w_j are uncorrelated; i.e.,

$$E\left(w_k w_j^*\right) = 0 \quad , \quad k \neq j \cdot \tag{D.52}$$

When $\{s_k\}$ is a stationary random process, the variance of w_k has a constant value; i.e.,

$$\omega \overset{\Delta}{=} E\{|w_k|^2\} \quad , \quad \text{arbitrary } k \cdot \tag{D.53}$$

The parameter ω is often referred to as the _prediction error._

An autoregressive process may be defined to be any (zero-mean) random process such that the estimate $\hat{s}_{k|k-1}$ is a fixed linear combination of a _finite_ number of past observations, i.e.,

$$w_k = s_k - (a_1 s_{k-1} + \ldots + a_L s_{k-L}) \cdot \tag{D.54}$$

Thus, an Lth order autoregressive (AR) process satisfies the Lth order linear difference equation

$$s_k = a_1 s_{k-1} + a_2 s_{k-2} + \ldots + a_L s_{k-L} + w_k . \tag{D.55}$$

The innovations of an AR process may be generated by applying the AR process to the input of a linear time invariant (LTI) filter; from Eq. (D.54), we conclude that this filter is of the (finite) moving average type with z-transform

$$A_L(z) = 1 - (a_1 z^{-1} + a_2 z^{-2} + \ldots + a_L z^{-L}) \tag{D.56}$$
$$= z^{-L} P_L(z)$$

where $P_L(z)$ is an Lth order polynomial with a unity leading coefficient. If all of the zeroes (roots) of $P_L(z)$ lie inside the unit circle, then $A(z)$ has a stable inverse $A^{-1}(z)$ (i.e., all the poles of $A^{-1}(z)$ lie inside the unit circle). Under this important __stability__ assumption, the AR process defined by the difference equation (D.55) is statistically __stationary__, and its power spectral density is obviously

$$\chi(\alpha) = |A^{-1}(e^{i\alpha})|^2 \omega_L \tag{D.57}$$
$$= \frac{\omega_L}{|P_L(e^{i\alpha})|^2} .$$

A standard problem in AR analysis is to generate the autoregressive coefficients $a_1, \ldots, a_L$ of a process from its known correlation lags. Quite often, the prediction error ω_L is also of interest. Note that when dealing with AR processes, the prediction error is given a subscript corresponding to the order of the process.

The relationship between the autoregressive coefficients and the correlation lags

$$c_\ell = E\{s_k s_{k-\ell}^*\} \quad ; \quad \text{k arbitrary} \tag{D.58}$$

is easily obtained by multiplying all the terms in (D.54) by s_{k-n}^* and taking

the expected value of the result. By the orthogonality principle, $E\{w_k\, s_{k-n}^*\}$ = 0 for n > 0, and we obtain

$$c_n - a_1\, c_{n-1} - a_2\, c_{n-2} - \cdots - a_L\, c_{n-L} = 0 \quad ; \quad n > 0 \; . \quad (D.59)$$

The first L of these equations may be arranged in the <u>normal</u> Yule-Walker form

$$
\begin{bmatrix}
c_0 & c_{-1} & \cdots & c_{-L+1} \\
c_1 & c_0 & & \vdots \\
\vdots & \vdots & \ddots & \vdots \\
\vdots & \vdots & \ddots & c_{-1} \\
\vdots & \vdots & \ddots & \vdots \\
c_{L-1} & \cdots & & c_0
\end{bmatrix}
\begin{bmatrix}
a_1 \\ \cdot \\ \cdot \\ \cdot \\ \cdot \\ a_L
\end{bmatrix}
=
\begin{bmatrix}
c_1 \\ \cdot \\ \cdot \\ \cdot \\ \cdot \\ c_L
\end{bmatrix}
\qquad (D.60)
$$

For n=0, we obtain an equation similar to Eq. (D.59), i.e.,

$$c_0 - a_1 c_{-1} - a_2 c_{-2} \cdots - d_L\, c_{-L} = E\{w_k\, s_k^*\} \; . \qquad (D.61)$$

Multiplying Eq. (D.54) by w_k^*, taking expected values, and invoking the orthogonality principle (again) shows that the right-hand side of Eq. (D.61) is real, i.e.,

$$E(w_k^* s_k) = E|w_k|^2 \; . \qquad (D.62)$$
$$= \omega_L$$

Substituting Eq. (D.62) in Eq. (D.61), and appending the first L equations from (D.59), we obtain the linear prediction equations

$$
\begin{bmatrix}
c_0 & c_{-1} & \cdots & c_{-L} \\
c_1 & & \ddots & \vdots \\
\vdots & & \ddots & c_{-1} \\
c_L & \cdots & c_1 & c_0
\end{bmatrix}
\begin{bmatrix}
1 \\ -a_1 \\ \vdots \\ -a_L
\end{bmatrix}
=
\omega_L
\begin{bmatrix}
1 \\ 0 \\ \vdots \\ 0
\end{bmatrix}
\qquad (D.63)
$$

Of course, the Yule-Walker <u>normal</u> equations and the linear prediciton equations are (essentially) equivalent; the choice between solving Eq. (D.60) and Eq. (D.3) 's often a matter of personal taste. In [95], Durbin presented a recursive procedure that solves the normal equations twice as fast as Levinson's original algorithm, which permits a general right-hand side for Eq. (D.60). In the next section, we derive the equivalent of Durbin's algorithm for the linear prediction equations in Eq. (D.63). We always assume that the Toeplitz matrix in Eq. (D.63) is Hermitian (recall $c_{-n} = c_n^*$) and at least positive semi-definite. In general, we refer to a positive semi-definite Hermitian Toeplitz matrix as a <u>correlation</u> matrix.

Linear Prediction

Consider a <u>positive-definite</u> correlation matrix in the standard (partitioned) form given below

$$
\underline{C}_{n+1} = \begin{bmatrix} c_0 & c_1^* & \cdot & \cdot & \cdot & c_n^* \\ c_1 & & \cdot & \cdot & & c_1^* \\ \vdots & & & \cdot & & \\ c_n & \cdot & \cdot & \cdot & c_1 & c_0 \end{bmatrix}
$$

$$
= \left[\begin{array}{c|c} c_0 & \underline{c}_{-n}^H \\ \hline \underline{c}_n & \underline{C}_{-n} \end{array} \right] \tag{D.64}
$$

where

$$
\underline{c}_n \triangleq (c_1 \ \ldots \ c_n)^T \tag{D.65}
$$

is an n-dimensional vector of the correlation lags, and $\underline{C}_n$ is the indicated $n \times n$ principal submatrix of $\underline{C}_{n+1}$. The nth order linear prediction (LP) problem is to solve

$$
\underline{C}_{n+1}
\begin{bmatrix}
1 \\
-a_1 \\
\cdot \\
\cdot \\
\cdot \\
-a_n
\end{bmatrix}
= \omega_n
\begin{bmatrix}
1 \\
0 \\
\cdot \\
\cdot \\
\cdot \\
0
\end{bmatrix}
\tag{D.66}
$$

for the prediction error ω_n and the n (complex) LP coefficients $\{a_1 \ldots a_n\}$. In order to discuss (D.66), it is convenient to introduce the $(n+1)$-dimensional vector of _error filter weights_

$$
\underline{w}_{n+1} \triangleq (1 \; -a_1 \; \ldots \; -a_n)^T
\tag{D.67a}
$$

as well as the unit vector

$$
\underline{\nu}_{n+1} \triangleq (1 \; 0 \; \ldots \; 0)^T
\tag{D.67b}
$$

of the same dimension. Thus, the problem in Eq. (D.66) may be put in the form

$$
\underline{C}_{n+1} \; \underline{w}_{n+1} - \omega_n \; \underline{\nu}_{n+1} = 0
\tag{D.68a}
$$

$$
\underline{\nu}_{n+1}^T \; \underline{w}_{n+1} = 0 \; .
\tag{D.68b}
$$

We first show that the solution to the $(n-1)^{st}$ order problem exists and is unique. Naturally, the same argument applies to the nth order case as well. Since $\underline{C}_n$ is (also) non-singular, $\omega_{n-1} = 0$ cannot be a solution of

$$
\underline{C}_n \; \underline{w}_n = \omega_{n-1} \; \underline{\nu}_n \quad .
\tag{D.69}
$$

Consequently, we may rearrange Eq. (D.69) to obtain

$$\omega_{n-1}^{-1} \, \underline{w}_n = \underline{C}_n^{-1} \underline{v}_n .$$

(D.70)

Thus, the $(n-1)^{st}$ order error filter weights are specified to within a scale factor by the first column of the inverse of the correlation matrix. Since the leading (first) element of $\underline{w}_n$ is constrained to be unity (i.e., $\underline{v}_n^T \underline{w}_n = 1$), it follows directly from Eq. (D.70) that the $(n-1)^{st}$ linear prediction error is

$$
\begin{aligned}
\omega_{n-1}^{-1} &= \underline{v}_n^T \underline{C}_n^{-1} \underline{v}_n \\
&= \frac{\text{cofactor of the upper left element of } \underline{C}_n}{|\underline{C}_n|} .
\end{aligned}
$$

(D.71)

The 'ratio of determinants' formula

$$\omega_n = \frac{|\underline{C}_{n+1}|}{|\underline{C}_n|}$$

(D.72)

follows from Eq. (D.71) and the observation, from Eq. (D.64), that $|\underline{C}_n|$ is the cofactor of the upper left element of the correlation matrix $\underline{C}_{n+1}$. Obviously, the result in Eq. (D.72) remains valid even when $\underline{C}_{n+1}$ is a general (non-singular) covariance matrix. However, the results obtained subsequently apply only to correlation matrices.

Levinson Recursion

The special case of the Levinson recursion [96] that solves the Yule-Walker normal equations most efficiently is usually credited to Durbin. It is currently more fashionable to solve (D.66) in order to obtain the prediction error (explicitly) along with the autoregressive coefficients. Burg [62] refers to the linear prediction algorithm presented below as the 'modern' Levinson recursion. Our derivation relies on the algebraic properties of the exchange operator introduced in Chapter II. The essential feature of our

argument is that a Toeplitz Hermitian matrix is always equal to its exchange (i.e., correlation matrices are harmonic).

We assume that the $(n-1)^{st}$ prediction error $\omega_{n-1} > 0$ is known and that the corresponding _error filter_ weights are available in the form of the n-dimensional vector $\underline{w}_n$ with a leading element of unity. We seek the solution of Eq. (D.61) in the form

$$\underline{w}_{n+1} = \underline{w}_{n+} - k_n \underline{w}_{n+}^\dagger \tag{D.73}$$

where $\underline{x}_+$ is a convenient notation for the 'one-step zero extension' of $\underline{x}$, i.e.,

$$\underline{x}_+ \overset{\Delta}{=} \frac{\underline{x}}{0} . \tag{D.74}$$

For convenience, the exchange of $\underline{w}_{n+}$ is written as $\underline{w}_{n+}^\dagger$ without parentheses[*]. Clearly, the form of the proposed solution in Eq. (D.73) preserves Eq. (D.68b). The nth (complex) reflection coefficient k_n is to be determined, together with $\omega_n > 0$, so that Eq. (D.73) solves Eq. (D.68a).

We first substitute Eq. (D.73) in Eq. (D.68a) and arrange the result as

$$\underline{C}_{n+1} \underline{w}_{n+} - \omega_n \underline{\nu}_{n+1} = k_n \underline{C}_{n+1} \underline{w}_{n+}^\dagger . \tag{D.75}$$

We now calculate

$$\underline{C}_{n+1} \underline{w}_{n+}^\dagger = \underline{C}_{n+1} \frac{0}{\underline{w}_n^\dagger} \tag{D.76}$$

---

[*]i.e., we always intend that $\underline{x}_+^\dagger$ be interpreted as $(\underline{x}_+)^\dagger$.

$$= \frac{\underline{c}_n^H \quad \underline{w}_n^\dagger}{\underline{c}_n^H \quad \underline{w}_n^\dagger}$$

using the partition of $\underline{C}_{n+1}$ in Eq. (D.64). By hypothesis, $\underline{w}_n$ solves Eq. (D.69), and it follows[**] that

$$\underline{C}_n \, \underline{w}_n^\dagger = \omega_{n-1} \, \underline{v}_n^\dagger .$$

Substituting this identity in Eq. (D.76) and introducing

$$e_n \triangleq (\underline{c}_n^{\,H} \, \underline{w}_n^{\,\dagger})^* \tag{D.77}$$

$$= \underline{c}_n^{\,\dagger H} \, \underline{w}_n$$

leads to

$$\underline{C}_{n+1} \, \underline{w}_{n+}^\dagger = \frac{e_n^*}{\omega_{n-1} \, \underline{v}_n^\dagger}$$

$$= e_n^* \, \underline{v}_{n+1} + \omega_{n-1} \, \underline{v}_{n+1}^\dagger . \tag{D.78}$$

Exchanging Eq. (D.78) yields a similar equation, i.e.,

$$\underline{C}_{n+1} \, \underline{w}_{n+} = e_n \, \underline{v}_{n+1}^\dagger + \omega_{n-1} \, \underline{v}_{n+1} . \tag{D.79}$$

Substituting both Eq. (D.78) and Eq. (D.79) directly into Eq. (D.75) leaves

$$(\omega_{n-1} - \omega_n) \, \underline{v}_{n+1} + e_n \, \underline{v}_{n+1}^\dagger = k_n \, e_n^* \, \underline{v}_{n+1} + k_n \, \omega_{n-1} \, \underline{v}_{n+1}^\dagger .$$

Identifying coefficients, we have only the trivial system

$$\omega_{n-1} - \omega_n = k_n \, e_n^* \tag{D.80a}$$

[**]Note that ω_{n-1}, $\underline{C}_n$, and $\underline{C}_{n+1}$ are all harmonic.

$$e_n = k_n \, \omega_{n-1} \tag{D.80b}$$

left to solve. Substituting Eq. (D.80b) in Eq. (D.80a) yields the well-known recursive relationship for the prediction errors

$$\omega_n = (1 - |k_n|^2) \, \omega_{n-1} \tag{D.81}$$

where the nth reflection coefficient k_n is

$$k_n = e_n / \omega_{n-1} \ . \tag{D.82}$$

The complex parameter e_n is easily calculated from (D.77) using the correlation coefficients (i.e., $\underline{c}_n$) and the $(n-1)^{st}$ order LP coefficients (i.e., $\underline{w}_n$). Written explicitly in terms of the $(n-1)^{st}$ order LP coefficients, we have

$$e_n = c_n - \sum_{m=1}^{n-1} a_m \, c_{n-m} \ . \tag{D.83}$$

Thus, e_n is sometimes interpreted as the "error in predicting the nth lag".

To show that the LP filter is stable, (i.e., $\omega_n > 0$), we find it convenient to express ω_{n-1} and e_n in terms of $\underline{w}_{n+}$. Pre-multiplying both sides of Eq. (D.79) by $\underline{w}_{n+}^{H}$ yields

$$\omega_{n-1} = \underline{w}_{n+}^{H} \ \underline{C}_{n+1} \ \underline{w}_{n+} \ . \tag{D.84}$$

Similarly, pre-multiplying Eq. (D.79) by $\underline{w}_{n+}^{\dagger H}$ gives

$$e_n = \underline{w}_{n+}^{+H} \ \underline{C}_{n+1} \ \underline{w}_{n+} \ . \tag{D.85}$$

Introducing the inner product

$$(\underline{x}, \underline{y})_{n+1} \triangleq \underline{x}^{H} \ \underline{C}_{n+1} \ \underline{y} \tag{D.86}$$

D-22

in Eqs. (D.84) and (D.85) leads to

$$k_n = \frac{(\underline{w}_{n+}^{\dagger}, \underline{w}_{n+})}{(\underline{w}_{n+}, \underline{w}_{n+})} \tag{D.87}$$

for the nth reflection coefficient defined in Eq. (D.82).

Since $\underline{C}_{n+1}$ is harmonic, we also have from Eq. (D.86)

$$\omega_{n-1} = (\underline{w}_{n+}^{\dagger}, \underline{w}_{n+}^{\dagger}) = (\underline{w}_{n+}, \underline{w}_{n+}) \ . \tag{D.88}$$

Thus, the Schwartz inequality guarantees $|k_n| <$ with equality iff $\underline{w}_{n+}^{\dagger} = c\ \underline{w}_{n+}$ for some non-zero (complex) c. However, $\underline{w}_{n+}$ and $\underline{w}_{n+}^{\dagger}$ are linearly independent by virtue of their zero-extensions, i.e.,

$$\underline{w}_{n+} = \begin{bmatrix} \underline{w}_n \\ 0 \end{bmatrix} \quad \Rightarrow \quad \underline{w}_{n+}^{\dagger} = \begin{bmatrix} 0 \\ \underline{w}_n^{\dagger} \end{bmatrix}.$$

Since the leading coefficient of $\underline{w}_n$ is unity (by hypothesis), we must conclude that the magnitude of the nth reflection coefficient is strictly less than unity, i.e.,

$$|k_n| < 1 \ . \tag{D.89}$$

Moreover, the recursive relationship (now) established in Eq. (D.73) guarantees that Eq. (D.89) will hold for each succeeding value of n. The recursion is started with $\omega_0 = c_0$, $\underline{w}_{0+} = (1\ \ 0)^T$ and hence Eq. (D.89) holds for every positive integer n.

Stability of LP Filters

The nth order filter $A_n(z)$ derived from the error filter weight $\underline{w}_{n+1}$ is shown below to be minimum phase and hence stable. The only assumptions needed for this argument are that $\underline{w}_n$ satisfies the recursive relationship in Eq. (D.73),

$$\underline{w}_{n+1} = \underline{w}_{n+} - k_n\ \underline{w}_{n+}^{\dagger}$$

for any set of reflection coefficients satisfying

$$|k_n| < 1$$

for all $n \geq 1$, where

$$\underline{w}_{0+} \triangleq \binom{1}{0}$$

and, for $n > 1$,

$$\underline{w}_{n+} \triangleq \begin{bmatrix} \underline{w}_n \\ \hline 0 \end{bmatrix} \quad .$$

We introduce the polynomial operator

$$\underline{p}^T_{n+1}(z) \triangleq (1 \; z \; \ldots \; z^n) \tag{D.90}$$

and define the nth order error filter (in the usual fashion) as the z-transform of the error filter weights, i.e.,

$$A_n(z) = \underline{p}^T_{n+1}(z^{-1}) \; \underline{w}_{n+1} \quad . \tag{D.91}$$

Actually, the stability argument is simplified somewhat by working with the (equivalent) error filter polynomials. Thus, the error polynomial is defined as

$$P_n(z) \triangleq \underline{p}^T_{n+1}(z) \; \underline{w}_{n+1} \quad . \tag{D.92}$$

The nth reciprocal error polynomial is defined to be

$$P^\dagger_n(z) \triangleq \underline{p}^T_{n+1}(z) \; \underline{w}^\dagger_{n+1} \tag{D.93}$$

and may be obtained from $P_n($) by 'exchanging' polynomial coefficients, i.e.,

$$P_n(z) = p_n z^n = \ldots + p_1 z + p_0 \quad ,$$

$$\Rightarrow P_n^\dagger(z) = p_0^* z^n + \ldots + p_{n-1}^* z + p_n^*$$

Szego [90] defines the reciprocal polynomial in the same way, but uses the notation $P^*(z)$. We prefer to use $P^*(z)$ to refer to the __conjugate__ polynomial

$$P^*(z) = p_n^* z^n + \ldots + p_1^* z + p_0^* \; .$$

Note that the conjugate of $P(z)$ must be written as

$$[P(z)]^* = P^*(z^*)$$

in order to be consistent with the notation introduced above.

A polynomial and its reciprocal (polynomial) have certain interesting analytic properties. For example, if z_1 is a root of $P(z)$, then $1/z_1^*$ is a root of $P^\dagger(z)$. Thus, $P(z)$ and $P^\dagger(z)$ exhibit a type of symmetry with respect to the unit circle. Moreover, it follows directly from the definition in Eq. (D.93) that

$$|P_n(e^{i\phi})| = |P_n^\dagger(e^{i\phi})| \; . \tag{D.94}$$

With our notation established, we combine Eq. (D.92) with Eq. (D.73) to obtain the polynomial recursion

$$
\begin{aligned}
P_n(z) &= P_{n+1}^T(z) \, [\underline{w}_{n+} - k_n \, \underline{w}_{n+}^\dagger] \\
&= P_{n-1}(z) - k_n \, z \, P_{n-1}^\dagger(z)
\end{aligned}
\tag{D.95}
$$

where the second term stems from the representation

$$P_{n+1}^T(z) = z \left(z^{-1} \; p_n^T(z) \right) \; .$$

We may now apply a fundamental result from the theory of analytic function to show that the zeroes of $P_n(z)$ are always outside the unit circle.

Rouche's Theorem [97]

Let $f(z)$ and $g(z)$ be analytic inside and on the unit circle C with $|g(z)| < |f(z)|$ on C. Then $f(z)$ and $f(z) + g(z)$ have the same number of zeroes inside C.

The previous assertion concerning the zeroes of $P_n(z)$ can now be established by induction; i.e., we assume that $P_{n-1}(z)$ has no zeroes inside or on the unit circle and show that $P_n(z)$ enjoys the same property. Obviously, the hypothesis is true for $P_0(z)$. We identify the polynomials $P_{n-1}(z)$ and $-k_n z P_{n-1}^\dagger(z)$ with $f(z)$ and $g(z)$ in Rouche's theorem. Since $|k_n| < 1$, Eq. (D.94) guarantees that Rouche's condition is met. Consequently, we conclude from Eq. (D.95) that $P_n(z)$ and $P_{n-1}(z)$ must have the same number of zeroes (i.e., none) inside the unit circle. Moreover, $P_n(z)$ cannot have any zeroes on the unit circle (either) without contradicting Eq. (D.89). By induction, all the zeroes of $P_n(z)$ lie outside the unit circle (for any positive integer n).

Since the z transform of the nth error filter weight is related to the nth error polynomial by

$$A_n(z) = P_n(z^{-1}) \qquad (D.96)$$

the zeroes of $A_n(z)$ must lie inside the unit circle (i.e., z_j is a zero of A_n iff z_j^{-1} is a zero of P_n).

Since $P_n(z^{-1}) = z^{-n} P_n^{\dagger *}(z)$, it follows that all the poles of $A_n(z)$ occur at the origin. A filter with all its poles and zeroes inside the unit circle is said to be minimum phase [74]. Therefore, any filter $A_n(z)$ derived (recursively) from weights satisfying the conditions specified in this section is minimum phase. Moreover, the inverse filter $A_n^{-1}(z)$ is also minimum phase (and hence stable).

Information Rate and the Prediction Error

For a stationary process with correlation lags $\{c_0, c_1, \ldots\}$, the limiting information rate (per sample) is

$$\gamma_\infty \stackrel{\Delta}{=} \lim_{n \to \infty} \gamma_n$$

where the average information rate for n samples (D.14) is written as

$$\gamma_n = \frac{1}{n} \ell_n |\underline{C}_n| \tag{D.97}$$

in the case of a stationary signal. The matrix $\underline{C}_n$ in (D.97) is specified in (D.64). The nth order prediction error for $\underline{C}_n$ is determined by (D.66).

It follows from (D.81) and (D.89) that the sequence of prediction errors satisfies

$$0 < \omega_n < \omega_{n-1} \tag{D.98}$$

and, consequently, a (unique) limiting prediction error

$$\omega_\infty \stackrel{\Delta}{=} \lim_{n \to \infty} \omega_n \tag{D.99}$$

is guaranteed to exist. Our goal here is to relate the information rate γ_∞ to the limiting prediction error ω_∞.

The following theorem is a standard exercise in many calculus (e.g., [98]) and real analysis (e.g., [99]) textbooks. The proof is not obvious, so is given below for the sake of completeness.

Theorem: Let $\{\Delta_n\}$ be any sequence of positive numbers. If $\lim \dfrac{\Delta_{n+1}}{\Delta_n} = \omega \ge 0$, then

$$\lim \Delta_n^{1/n} = \omega \quad .$$

Proof: The trick is to first write the identity

$$\frac{\Delta_{n+1}}{\Delta_n} = \frac{\Delta_{n+1}}{\Delta_n} \frac{\Delta_n}{\Delta_{n-1}} \cdots \frac{\Delta_{N+1}}{\Delta_N} \frac{\Delta_N}{\Delta_n}$$

for any $n > N$. Given $\epsilon > 0$, there exists a positive integer N such that

$$\omega - \varepsilon < \frac{\Delta_{n+1}}{\Delta_n} < \omega + \varepsilon$$

for any $n > N$.

Consequently, we obtain

$$(\omega - \varepsilon)^{n+1-N} \frac{\Delta_N}{\Delta_n} < \frac{\Delta_{n+1}}{\Delta_n} < (\omega + \varepsilon)^{n+1-N} \frac{\Delta_N}{\Delta_n} .$$

Multiplying these inequalities through by (positive) Δ_n yields

$$\alpha (\omega - \varepsilon)^{n+1} < \Delta_{n+1} < \beta(\omega - \varepsilon)^{n+1}$$

where we have put $\alpha = \Delta_N/(\omega - \varepsilon)^N$ and $\beta = \Delta_N/(\omega + \varepsilon)^N$.

It follows that

$$\alpha^{1/n}(\omega - \varepsilon) < \Delta_n^{1/n} < \beta^{1/n}(\omega + \varepsilon) .$$

Since $\lim \alpha^{1/n} = \lim \beta^{1/n} = 1$ (assuming positive α), taking limits gives

$$\omega - \varepsilon < \lim \Delta_n^{1/n} < \omega + \varepsilon .$$

Since ε can be chosen arbitrarily small, the theorem follows directly for positive ω. If $\omega = 0$, $\Delta_n^{1/n}$ is bounded from below by 0 and above by ε, thus proving the theorem for $\omega = 0$.

This theorem establishes the well-known relationship between the (Wiener) linear prediction error ω_∞ and the (Shannon) Gaussian information rate γ_∞. Setting $\Delta_n = |\underline{C}_{-n}|$, it follows from (D.97) that

$$\omega_n = \frac{\Delta_{n+1}}{\Delta_n} \quad ,$$

and from (D.97) we obtain

$$\gamma_n = \ln \Delta_n^{1/n} .$$

Taking limits and applying the above theorem gives

$$\gamma_\infty = \ln \omega_\infty \ .$$

APPENDIX E

A VECTOR SPACE FORMULATION OF THE SUM OF
SINUSOIDS/PLANE WAVES PROBLEM*

A unifying geometrical interpretion based on least-squares concepts is given to the problem of measuring the parameters of a finitely sampled signal composed of a sum of complex sinusoids (e.g., plane waves). Special attention is given to the problem of order determination, i.e., determining the number of distinct sinusoids present.[†]

The problem is seen to have two important facets, one of approximation and one of estimation. Every signal of the assumed form has a precisely defined order; however, different order signals are not equally distinguishable. Every signal of a given order is characterized not only by its order but also by how closely a lower order signal can approximate it. A vector norm definition seems appropriate for quantifying the approximation accuracy.

The minimum approximation distance by lower order models, a purely deterministic parameter associated with each signal, is shown by Cramer-Rao bounding arguments to be an appropriate measure of "signal strength" for the problem of estimating the true order when the signal is observed in white Gaussian noise. The definition of effective signal-to-noise ratio under this criterion

---

*This self-contained appendix is drawn from an internal Lincoln Laboratory memorandum by S. Dolinar. The notation used here is different from that of the body of the report, but we have indicated the correspondences at appropriate points.

[†]This is not the same as determining the order of the best-fitting AR or ARMA model; rather it corresponds to counting the number of peaks in the AR(MA) estimated spectrum.

is shown to reduce to the usual definition in terms of signal component ampli-
tudes whenever the signal frequencies are widely spaced. For closely spaced
frequencies the effective signal-to-noise ratio is found to depend on other
parameters as well, most notably the frequency separation and relative phase.

An order determination algorithm based on a maximum likelihood estimate
of the minimum approximation distance is suggested. Bounds on the error
performance of the order estimator are calculated, and the behavior of the
algorithm is further examined by applying it to computer-generated synthetic
data. It is shown that accurate order determination of a closely spaced two-
frequency signal is possible for signal amplitudes significantly lower than
those that have so far been needed for accurate AR or ARMA estimates. In
fact, it is argued that an unrestricted ARMA model is probably too general to
permit optimum order resolution.

The problem of maximum likelihood estimation of the remaining signal
parameters is also seen to have a useful interpretation in the vector space
formulation. The signal component amplitudes and phases are determined by
simply projecting the signal into a subspace spanned by vectors associated
with each signal component frequency. The signal frequencies are determined
by maximizing this projection.

1. Problem Formulation

We consider the problem of detecting several complex sinusoidal signals
or plane waves in white Gaussian noise. The observed data consist of N
equally spaced samples $\{x_k\}$, where

$$x_k = s_k + n_k \qquad\qquad k = 1, \ldots, N \qquad\qquad (1)$$

$$s_k = \sum_{i=1}^{I} \alpha_i \, \beta_i^{\left(k - \frac{N+1}{2}\right)} \qquad k = 1, \ldots, N \qquad\qquad (2)$$

$$\alpha_i = A_i \, e^{j\phi_i}, \qquad\qquad i = 1, 2, \ldots, I \ \text{(unknown)} \qquad (3)$$

$$\beta_i = e^{j\omega_i \, \Delta T}, \qquad\qquad i = 1, 2, \ldots, I \ \text{(unknown)} \qquad (4)$$

$$N\Delta T = 1 \, ^* \qquad\qquad\qquad\qquad\qquad\qquad\qquad\qquad (5)$$

and $\{n_k\}$ is a complex white Gaussian noise process, with

$$E \, n_k = E \, n_k \, n_\ell = 0 \qquad\qquad\qquad\qquad\qquad (6)$$

$$E \, n_k^* \, n_\ell = \sigma^2 \, \delta_{k\ell} \qquad . \qquad\qquad\qquad\qquad (7)$$

To simplify the analysis we introduce vector notation, defining

---

*This condition implies that the frequencies $\{\omega_i\}$ are measured in units of $2\pi \times$ beamwidths, where one beamwidth is the frequency (in Hz) required to produce a phase differential of $2\pi(N-1)/N$ between the 1st sample and the Nth sample.

$$\underline{s} = [s_1\ s_2\ \cdots\ s_N]^T \tag{8}$$

$$\underline{\beta}_i = \left[\beta_i^{-\frac{N-1}{2}}\ \cdots\ \beta_i^{\frac{N-1}{2}}\right]^T = \left[e^{-j\omega_i\Delta T\frac{N-1}{2}}\ \cdots\ e^{j\omega_i\Delta T\frac{N-1}{2}}\right]^T \tag{9}$$

$$\underline{n} = [n_1\ n_2\ \cdots\ n_N]^T \tag{10}$$

$$\underline{x} = [x_1\ x_2\ \cdots\ x_N]^T \quad . \tag{11}$$

Then equations (1) and (2) can be rewritten

$$\underline{x} = \underline{s} + \underline{n} \tag{12}$$

$$\underline{s} = \sum_{i=1}^{I} \alpha_i\ \underline{\beta}_i \quad . \tag{13}$$

The above signal model is related to the signal model introduced in Chapter II as follows:

This Appendix	Chapter II
$\alpha_i\ (= A_i e^{j\phi_i})$	p_i
β_i	α_i
$w_i\Delta T$	$2\pi\frac{\xi}{\lambda}\sin\theta_i$
$\underline{\beta}_i$	$\underline{v}\ (\alpha_i)$
$\underline{x}$	$\underline{r}$

$$\underline{s} \qquad\qquad \underline{s}$$

The vectors defined above are assumed to be elements of an inner product space V, with inner product function and vector norm defined by

$$(\underline{u}, \underline{v}) = \sum_{k=1}^{N} w_k u_k^* v_k \qquad\qquad (14)$$

$$|\underline{v}| = \sqrt{(\underline{v}, \underline{v})} \qquad\qquad (15)$$

where $\{w_k\}$ is a real window sequence satisfying

$$w_k > 0 \qquad\qquad k = 1, \ldots, N \qquad\qquad (16)$$

and

$$\sum_{k=1}^{N} w_k = 1 \qquad . \qquad\qquad (17)$$

We will normally assume that $\{w_k\}$ is uniform,

$$w_k = \frac{1}{N} \qquad\qquad k = 1, \ldots, N \qquad\qquad (18)$$

but many of our results are not dependent on this assumption.

2. Lth Order Approximations to a Known Signal

We consider Lth order approximations to the signal vector $\underline{s}$ in the form

$$\underline{\tilde{s}}^{(L)} = \sum_{i=1}^{L} \tilde{\alpha}_i \underline{\tilde{\beta}}_i , \qquad \underline{\tilde{\beta}}_i \, \epsilon \, B , \; \tilde{\alpha}_i \text{ complex numbers} \qquad (19)$$

where the set of admissible representation vectors consists of all $\underline{\beta}$ of the form (9), i.e.,

$$B = \left\{ \underline{\beta} = [\beta_1 \; \cdots \; \beta_N]^T : \beta_k = e^{j\omega\Delta T(k - \frac{N+1}{2})} \text{ for some real } \omega \right\} . \qquad (20)$$

The accuracy of the Lth order approximation is measured by the distance $\tilde{\delta}^{(L)}$ between $\underline{\tilde{s}}^{(L)}$ and $\underline{s}$.

$$\tilde{\delta}^{(L)} = \| \underline{s} - \underline{\tilde{s}}^{(L)} \| . \qquad (21)$$

The most accurate Lth order approximation is the one which minimizes $\tilde{\delta}^{(L)}$; the resulting minimum approximation distance is denoted $\tilde{\delta}^{(L)}_{min}$.

Part of the minimization problem leading to the evaluation of the most accurate Lth order approximation is straightforward. For arbitrary fixed representation vectors $\{\underline{\tilde{\beta}}_i\}$, the optimum choice of complex amplitudes $\{\tilde{\alpha}_i\}$ produces an approximation vector $\underline{\tilde{s}}^{(L)}$ equal to the projection of the true signal vector $\underline{s}$ into the subspace spanned by $\{\underline{\tilde{\beta}}_i\}$. The determination of the optimum $\{\tilde{\alpha}_i\}$ is thus reduced to the solution of a set of linear equations,

$$\sum_{j=1}^{L} (\underline{\tilde{\beta}}_i, \underline{\tilde{\beta}}_j) \, \tilde{\alpha}_j = (\underline{\tilde{\beta}}_i, \underline{s}) \qquad i = 1, 2, \ldots, L . \qquad (22)$$

For $\{\tilde{\alpha}_i\}$ given in terms of $\{\underline{\tilde{\beta}}_i\}$ by (22), the approximation distance is evaluated as

$$[\tilde{\delta}^{(L)}]^2 = |\underline{s}|^2 - |P_L(\underline{\tilde{\beta}}_1, \ldots, \underline{\tilde{\beta}}_L)\,\underline{s}|^2 \tag{23}$$

where $P_L(\underline{\tilde{\beta}}_1, \ldots, \underline{\tilde{\beta}}_L)$ is the projection operator for the L-dimensional subspace spanned by $\underline{\tilde{\beta}}_1, \ldots, \underline{\tilde{\beta}}_L$. The optimum set of representation vectors $\{\underline{\tilde{\beta}}_i^0\}$ is the one which maximizes $|P_L(\underline{\tilde{\beta}}_1, \ldots, \underline{\tilde{\beta}}_L)\,\underline{s}|^2$ for admissible $\underline{\tilde{\beta}}_i \in B$.

The optimum approximation distance $\tilde{\delta}_{min}^{(L)}$ is a monotonically decreasing function of the approximation order L. For $L \geq I$ (I is the true order), the approximation can be made exact by taking $\underline{\tilde{\beta}}_i = \underline{\beta}_i$, $i = 1, \ldots, I$. For $L < I$, the non-zero approximation distance $\tilde{\delta}_{min}^{(L)}$ measures the extent to which the true I^{th} order signal can be distinguished from an Lth order model. For example, a second-order signal with two comparable amplitudes and widely spaced frequencies is characterized by $\tilde{\delta}_{min}^{(2)} = 0$ and $\tilde{\delta}_{min}^{(1)}$ large, whereas a second-order signal with closely spaced frequencies or one very small amplitude has $\tilde{\delta}_{min}^{(2)} = 0$ and $\tilde{\delta}_{min}^{(1)}$ nearly zero. For the latter example, a very small uncertainty in the observation of $\tilde{\delta}_{min}^{(L)}$, $L = 1, 2$, will prevent one from determining the true order.

3. Estimating the Lth Order Approximation Distance in the Presence of Noise

When the observation of the true signal is corrupted by noise, the optimum L^{th} order approximation can only be estimated. Limits on the achievable estimation accuracy can be inferred from the Cramer-Rao bound.

In particular, we are interested in estimating the Lth order approximation distance $\tilde{\delta}^{(L)}_{min}$ when the true signal is observed in additive white Gaussian noise, as in (1) - (13). For this noise model, the Cramer-Rao bound on the accuracy of any unbiased estimate $\hat{b}$ of an arbitrary unknown signal parameter b takes the form[*]

$$Var(\hat{b}) \geqslant \left[\frac{2N}{\sigma^2} \parallel \partial \underline{s}/\partial b \parallel^2 \right]^{-1} \quad . \qquad (24)$$

In order to calculate the derivative of $\underline{s}$ with respect to $\tilde{\delta}^{(L)}_{min}$, we express $\underline{s}$ in the form

$$\underline{s} = P_L(\tilde{\underline{\beta}}^O_1, \ldots, \tilde{\underline{\beta}}^O_L)\underline{s} + \tilde{\delta}^{(L)}_{min} \underline{s}^{\perp}(\tilde{\underline{\beta}}^O_1, \ldots, \tilde{\underline{\beta}}^O_L) \qquad (25)$$

where $\underline{s}^{\perp}(\tilde{\underline{\beta}}^O_1, \ldots, \tilde{\underline{\beta}}^O_L)$ is the optimum L^{th} order approximation error vector, normalized to unit norm.

$$\underline{s}^{\perp}(\tilde{\underline{\beta}}^O_1, \ldots, \tilde{\underline{\beta}}^O_L) = [\underline{s} - P_L(\tilde{\underline{\beta}}^O_1, \ldots, \tilde{\underline{\beta}}^O_L)\underline{s}]/\parallel \underline{s} - P_L(\tilde{\underline{\beta}}^O_1, \ldots, \tilde{\underline{\beta}}^O_L)\underline{s} \parallel \qquad (26)$$

Thus, if $\hat{\delta}^{(L)}$ is any unbiased estimate of $\tilde{\delta}^{(L)}_{min}$, it follows from (24) that

---

[*]The uniform window assumption (18) is made in writing this expression.

$$\text{Var}(\hat{\delta}^{(L)}) \geqslant \frac{\sigma^2}{2N} \quad . \tag{27}$$

This result helps to quantify the uncertainty in determining the true signal order I by estimating $\hat{\gamma}_{min}^{(L)}$. The two cases $\hat{\gamma}_{min}^{(L)} = 0$ (implying $I \leq L$) and $0 < \hat{\gamma}_{min}^{(L)} < \sqrt{\sigma^2/2N}$ (for which $I > L$) will be very difficult to distinguish when noise is present.

4. Estimating the Signal Order

A simple algorithm is suggested for estimating the true signal order. First, a sequence of Lth order approximation distance estimates $\hat{\delta}^{(L)}$, $L = 0, 1, 2, \ldots$, is generated from the observed data $\underline{x}$. These estimates should embody the monotonicity constraint $\hat{\delta}(0) \geqslant \hat{\delta}(1) \geqslant \hat{\delta}(2) \geqslant \ldots$ and it would be nice if the accuracy of the estimators approached the Cramer-Rao bound. Next, a suitable threshold ϵ is determined on the basis of the estimator variance and the desired resolvability. The estimate $\hat{I}$ of the true signal order I is given by

$$\hat{I} = \min \{L: \hat{\delta}^{(L)} < \epsilon\} \quad . \tag{28}$$

The estimate $\hat{I}$ may be either too high, too low, or exactly correct. The two incorrect types of estimates will be described in binary detection nomenclature as "false alarm" and "miss" events, respectively, occurring with probabilities $P_F^{(I)}$, $P_M^{(I)}$.

$$P_F^{(I)} = \Pr(\hat{I} > I) = \Pr|\hat{\delta}^{(I)} > \epsilon| \tag{29}$$

$$P_M^{(I)} = Pr \; (\hat{I} < I) = Pr|\hat{\delta}^{(I-1)} < \epsilon| \qquad (30)$$

The simple evaluation of $P_F^{(I)}$, $P_M^{(I)}$ follows from the assumed monotonicity of $\{\hat{\delta}^{(L)}\}$.

The "false alarm" and "miss" probabilities are more conveniently expressed in terms of the normalized zero mean, unit variance random variables

$$d_F^{(I)} = |\hat{\delta}^{(I)} - E \; \hat{\delta}^{(I)}| \; / \; \sqrt{Var(\hat{\delta}^{(I)})} \qquad (31)$$

$$d_M^{(I)} = |E \; \hat{\delta}^{(I-1)} - \hat{\delta}^{(I-1)}| \; / \; \sqrt{Var(\hat{\delta}^{(I-1)})} \qquad (32)$$

and the normalized thresholds

$$\theta_F^{(I)} = |\epsilon - E \; \hat{\delta}^{(I)}| \; / \; \sqrt{Var(\hat{\delta}^{(I)})} \qquad (33)$$

$$\theta_M^{(I)} = |E \; \hat{\delta}^{(I-1)} - \epsilon| \; / \; \sqrt{Var[\hat{\delta}^{(I-1)}]} \qquad . \qquad (34)$$

The probabilities $P_F^{(I)}$, $P_M^{(I)}$ are then given by

$$P_F^{(I)} = Pr \; |d_F^{(I)} > \theta_F^{(I)}| \qquad (35)$$

$$P_M^{(I)} = \Pr \left| d_M^{(I)} > \theta_M^{(I)} \right| \qquad . \qquad (36)$$

If the estimates $\hat{\delta}^{(I)}$, $\hat{\delta}^{(I-1)}$ are unbiased the expectations required in (31) – (34) are calculated as

$$E \; \hat{\delta}^{(I)} = 0 \qquad (37)$$

$$E \; \hat{\delta}^{(I-1)} = \tilde{\delta}_{min}^{(I-1)} \qquad (38)$$

and if the estimators achieve the Cramer-Rao accuracy bound, the variances are given by

$$Var\left| \hat{\delta}^{(I-1)} \right| = Var\left| \hat{\delta}^{(I)} \right| = \frac{\sigma^2}{2N} \qquad . \qquad (39)$$

Thus, the normalized error probability thresholds satisfy

$$\theta_F^{(I)} + \theta_M^{(I)} = \tilde{\delta}_{min}^{(I-1)} \sqrt{2N/\sigma^2} \qquad . \qquad (40)$$

The expression (40) implies a trade-off relation between the observation noise level and the minimum $(I-1)^{th}$-order approximation distance. For any fixed minimum performance levels $\theta_F^{(I)}$, $\theta_M^{(I)}$, the maximum acceptable noise level σ is proportional to the distance $\tilde{\delta}_{min}^{(I-1)}$ between the true I^{th}-order signal and the best $(I-1)^{th}$-order approximation to it. Thus, the minimum $(I-1)^{th}$-order approximation distance may be regarded as the effective "signal strength" for the order determination problem, and an effective signal-to-noise ratio may be defined by

$$SNR_{eff}^{(I)} = \frac{N}{\sigma^2} |\widetilde{\delta}_{min}^{(I-1)}|^2 \quad . \tag{41}$$

5. Evaluation of the Minimum L^{th} Order Approximation Distance

a. General Formulation

The approximation problem considered in Section 2 can be restated as follows: Given a signal vector

$$\underline{s} = \sum_{i=1}^{I} \alpha_i \underline{\beta}_i \quad , \qquad \underline{\beta}_i \in B \tag{42}$$

find L admissible representation vectors $\underline{\widetilde{\beta}}_1, \ldots, \underline{\widetilde{\beta}}_L$, $\underline{\widetilde{\beta}}_i \in B$, which maximize $\|P_L(\underline{\widetilde{\beta}}_1, \ldots, \underline{\widetilde{\beta}}_L) \underline{s}\|$, the norm of the projection of $\underline{s}$ into the subspace spanned by $\underline{\widetilde{\beta}}_1, \ldots, \underline{\widetilde{\beta}}_L$. We assume $L < I$ because the solution for $L > I$ is trivial.

In order to establish a connection between this geometrical problem and the physically meaningful frequency parameters $\{\omega_i\}$, $\{\widetilde{\omega}_i\}$ underlying the definitions of the signal component vectors $\{\underline{\beta}_i\}$ and the approximation vectors $\{\underline{\widetilde{\beta}}_i\}$, it is necessary to be more specific about the geometrical relationship among the vectors in B. The inner product between vectors $\underline{\beta}, \underline{\beta}' \in B$, where $\beta_k = e^{j\omega'\Delta T(k - \frac{N+1}{2})}$, is calculated as

$$(\underline{\beta}, \underline{\beta}') = W(\omega - \omega') \tag{43}$$

where $W(\cdot)$ is the discrete Fourier transform of the (centralized) window sequence $w_{k+(N+1)/2}$,

$$W(\omega) = e^{j\omega\Delta T\frac{N+1}{2}} \sum_{k=1}^{N} w_k e^{-j\omega\Delta Tk} .$$ (44)

The vectors $\underline{\beta} \in B$ all have unit norm

$$|\underline{\beta}|^2 = W(0) = \sum_{k=1}^{N} w_k = 1$$ (45)

and two vectors $\underline{\beta}, \underline{\beta}' \in B$ are approximately orthogonal if their frequency separation $|\omega - \omega'|$ is outside the effective bandwidth of the window sequence. For the uniform window sequence $(w_k = \frac{1}{N})$, $W(\cdot)$ is given by

$$W(\omega) = \frac{\sin\frac{\omega}{2}}{N\sin\frac{\omega}{2N}} \approx \text{sinc } \omega/2 \text{ if } N \text{ is large}$$ (46)

where

$$\text{sinc } x \equiv \frac{\sin x}{x} .$$ (47)

Since the window spectrum $W(\cdot)$ completely specifies the inner product structure of the set of admissible representation vectors B, it is possible to reformulate the approximation problem in terms of $W(\cdot)$. We write the Lth – order approximation distance as

$$|\tilde{\delta}^{(L)}|^2 = (\underline{s}, \underline{s} - P_L(\tilde{\underline{\beta}}_1, \ldots, \tilde{\underline{\beta}}_L)\underline{s})$$

$$= (\sum_{i=1}^{I} \alpha_i \underline{\beta}_i, \sum_{i=1}^{I} \alpha_i \underline{\beta}_i - \sum_{i=1}^{L} \tilde{\alpha}_i \tilde{\underline{\beta}}_i)$$

$$= \sum_{i=1}^{I} \sum_{j=1}^{I} \alpha_i^* \alpha_j (\underline{\beta}_i, \underline{\beta}_j) - \sum_{i=1}^{I} \sum_{j=1}^{L} \alpha_i^* \tilde{\alpha}_j (\underline{\beta}_i, \tilde{\underline{\beta}}_j) \quad .$$
$$(48)$$

Referring to (22) for the evaluation of the projection coefficients $\tilde{\alpha}_i$, we rewrite (48) in the form

$$|\tilde{\delta}^{(L)}|^2 = \underline{\alpha}^+ (W_{II} - W_{LI}^+ W_{LL}^{-1} W_{LI}) \underline{\alpha} \tag{49}$$

where

$$\underline{\alpha} = [\alpha_1, \ldots \alpha_L]^T \qquad \underline{\alpha}^+ = [\alpha_1^* \ldots \alpha_L^*] \tag{50}$$

and W_{II}, W_{LI}, W_{LL}, W_{LI}^+ are matrices (of dimension $I \times I$, $L \times I$, $L \times L$, $I \times L$, respectively) defined by

$$(W_{II})_{ij} = (\underline{\beta}_i, \underline{\beta}_j) = W(\omega_i - \omega_j) \tag{51}$$

$$(W_{LI})_{ij} = (\tilde{\underline{\beta}}_i, \underline{\beta}_j) = W(\tilde{\omega}_i - \omega_j) \tag{52}$$

$$(W_{LL})_{ij} = (\tilde{\underline{\beta}}_i, \tilde{\underline{\beta}}_j) = W(\tilde{\omega}_i - \tilde{\omega}_j) \tag{53}$$

$$(W_{LI}^+)_{ij} = (\underline{\beta}_i, \tilde{\underline{\beta}}_j) = W(\omega_i - \tilde{\omega}_j) = (W_{LI})_{ji}^* . \tag{54}$$

The first term in (49) equals $|\underline{s}|^2$ and the second term is an explicit evaluation of $|P_L(\tilde{\underline{\beta}}_1, \ldots, \tilde{\underline{\beta}}_L)\underline{s}|^2$, i.e.,

$$|\underline{s}|^2 = \underline{\alpha}^+ W_{II} \underline{\alpha} \tag{55}$$

$$|P_L(\underline{\beta}_1, \ldots, \underline{\beta}_L)\underline{s}|^2 = \underline{\alpha}^+ W_{LI}^+ W_{LL}^{-1} W_{LI} \underline{\alpha} . \tag{56}$$

b. Special Cases

There are some special cases for which an approximate solution to the approximation problem has been obtained.

(1) Widely Spaced Signal Frequencies

For this case we assume that the signal frequencies $\{\omega_i\}$ are separated by at least twice the effective (one-sided) bandwidth of the window spectrum $W(\omega)$. This implies that no approximation frequency $\tilde{\omega}_i$ can be close to two or more signal frequencies ω_j, $\omega_{j'}$, simultaneously. Specifically, we assume that for every approximation frequency $\tilde{\omega}_i$ there is a unique signal frequency index

$m(i)$ such that

$$(\widetilde{\underline{\beta}}_i, \underline{\beta}_j) = W(\widetilde{\omega}_i - \omega_j) \approx 0 \text{ for all } j \neq m(i) . \tag{57}$$

This implies that

$$(\underline{\beta}_i, P_L(\widetilde{\underline{\beta}}_1, \ldots, \widetilde{\underline{\beta}}_L) \underline{\beta}_j) = 0 \quad \text{if } i \neq j \tag{58}$$

and

$$(\underline{\beta}_i, P_L(\widetilde{\underline{\beta}}_1, \ldots, \widetilde{\underline{\beta}}_L)\beta_i) = 0 \quad \text{if } i \neq m(j) \text{ for any } j = 1, \ldots, L . \tag{59}$$

Therefore

$$\|P_L(\widetilde{\underline{\beta}}_1, \ldots, \widetilde{\beta}_L)\underline{s}\|^2 = (\underline{s}, P_L(\widetilde{\underline{\beta}}_1, \ldots, \widetilde{\underline{\beta}}_L)\underline{s})$$

$$= \sum_{i=1}^{I} \sum_{j=1}^{I} \alpha_i^* \alpha_j (\underline{\beta}_i, P_L(\widetilde{\underline{\beta}}_1, \ldots, \widetilde{\underline{\beta}}_L)\underline{\beta}_j)$$

$$= \sum_{i=1}^{I} |\alpha_i|^2 (\underline{\beta}_i, P_L(\widetilde{\underline{\beta}}_1, \ldots, \widetilde{\underline{\beta}}_L)\underline{\beta}_i)$$

$$= \sum_{i=1}^{L} |\alpha_{m(i)}|^2 (\underline{\beta}_{m(i)}, P_L(\widetilde{\underline{\beta}}_1, \ldots, \widetilde{\underline{\beta}}_L)\underline{\beta}_{m(i)})$$

$$< \sum_{i=1}^{L} |\alpha_{m(i)}|^2$$

$$< \max_{m(\cdot)} \sum_{i=1}^{L} |\alpha_{m(i)}|^2. \tag{60}$$

This upper bound can be achieved by an obvious choice of representation vectors,

$$\underline{\tilde{\beta}}_i = \underline{\hat{\beta}}_{m(i)} \qquad i = 1, \ldots, L \tag{61}$$

where $\hat{m}(\cdot)$ is the function which re-orders the signal component indices in descending amplitude order, i.e.,

$$|\hat{\alpha}_{m(1)}| > |\hat{\alpha}_{m(2)}| > \ldots > |\hat{\alpha}_{m(I)}| \; . \tag{62}$$

Thus, the $\{\underline{\tilde{\beta}}_i^{(L)}\}$ defined by (61) are optimum and the minimum approximation distance is calculated as

$$\tilde{\delta}_{min}^{(L)} = \sum_{i=L+1}^{I} |\hat{\alpha}_{m(i)}|^2 \; . \tag{63}$$

The effective signal-to-noise ratio for the order determination problem is given simply by

$$SNR_{eff}^{(I)} = \frac{N}{\sigma^2} |\hat{\alpha}_{m(I)}|^2 = \frac{N}{\sigma^2} \min_i |\alpha_i|^2 \equiv SNR_{amp} \tag{64}$$

i.e., the effective signal-to-noise ratio is equal to the array signal-to-noise ratio SNR_{amp} defined in terms of the _amplitude_ of the weakest component. Therefore, the exact signal order can be determined with high probability as long as the (amplitude) signal-to-noise ratio of the weakest component is high enough.

(ii) Closely Spaced Frequencies: First Order Approximation

We consider first-order ($L = 1$) approximations to a signal of arbitrary true order I. The results are most useful for the case $I = 2$ because, as shown earlier, the minimum $(I-1)^{th}$-order approximation distance $\delta_{min}^{(I-1)}$ is the key parameter for determining the resolvability of an Ith-order signal.

The signal frequencies are assumed to be reasonably closely spaced. Specifically, we assume that all frequency differences involved in the definitions (51) – (54) of W_{II}, W_{LI}, W_{LL} are small enough that the following approximation to the window spectrum $W(\omega)$ is valid:

$$W(\omega) \approx \cos \gamma \omega \qquad (65)$$

where

$$\gamma^2 = \frac{1}{12} (1 - \frac{1}{N^2}) \text{ for uniform } \{w_k\} \ . \qquad (66)$$

With this definition of γ the approximation (65) has the same Taylor series expansion as the exact uniform window spectrum (46), through third-order

terms. For $|\omega| \leq \pi$ (separations within half a beamwidth), the approximation is accurate within 3%.

For $L = 1$ the matrix W_{LL} is simply a scalar (equal to unity) and the matrix W_{LI} is a row vector. Thus, the ij^{th} component of the $I \times I$ matrix used in (49) to calculate the minimum approximation distance is given by

$$(W_{II} - W_{LI}^+ W_{LL}^{-1} W_{LI})_{ij} = W(\omega_i - \omega_j) - W(\tilde{\omega}_1 - \omega_i) W(\tilde{\omega}_1 - \omega_j)$$

$$\approx \cos[\gamma(\tilde{\omega}_1 - \omega_j) - \gamma(\tilde{\omega}_1 - \omega_i)] - \cos[\gamma(\tilde{\omega}_1 - \omega_i)] \cos[\gamma(\tilde{\omega}_1 - \omega_j)]$$
$$= \sin[\gamma(\tilde{\omega}_1 - \omega_j)] \sin[\gamma(\tilde{\omega}_1 - \omega_i)] . \tag{67}$$

Thus,

$$|\tilde{\delta}^{(1)}|^2 = \left| \sum_{i=1}^{I} \alpha_i \sin[\gamma(\tilde{\omega}_1 - \omega_i)] \right|^2$$

$$= \frac{1}{4} \left| \sum_{i=1}^{I} \alpha_i \left[e^{j\gamma(\tilde{\omega}_1 - \omega_i)} - e^{-j\gamma(\tilde{\omega}_1 - \omega_i)} \right] \right|^2$$

$$= \frac{1}{4} \left\{ \left| \sum_{i=1}^{I} \alpha_i e^{-j\gamma\omega_i} \right|^2 + \left| \sum_{i=1}^{I} \alpha_i e^{j\gamma\omega_i} \right|^2 \right.$$

$$\left. - 2\text{Re}\left[e^{j2\gamma\tilde{\omega}_1} \sum_{i=1}^{I} \alpha_i e^{-j\gamma\omega_i} \sum_{j=1}^{I} \alpha_j^* e^{-j\gamma\omega_j} \right] \right\} . \tag{68}$$

This expression is obviously minimized by selecting $\tilde{\omega}_1$ to give the last term zero phase. The resulting minimum approximation distance is

$$\tilde{\delta}_{min}^{(1)} = \frac{1}{2} \left| \left| \sum_{i=1}^{I} \alpha_i e^{j\gamma\omega_i} \right| - \left| \sum_{i=1}^{I} \alpha_i e^{-j\gamma\omega_i} \right| \right| . \tag{69}$$

(iii) Closely Spaced Frequencies: Second-Order Signal

We now consider the previous case specialized to $I = 2$ (second-order signal). The minimum first-order approximation distance is expressed as

$$\tilde{\delta}_{min}^{(1)} = \frac{1}{2} \max \left(|\alpha_1|, |\alpha_2| \right) \tag{70}$$

$$\times \left| \sqrt{1 + \rho^2 + 2\rho \cos(\phi + \gamma\Delta\omega)} - \sqrt{1 + \rho^2 + 2\rho \cos(\phi - \gamma\Delta\omega)} \right|$$

where

$$\rho e^{j\phi} \equiv \begin{array}{l} \alpha_2/\alpha_1 \text{ if } |\alpha_2| < |\alpha_1| \\ \alpha_1/\alpha_2 \text{ if } |\alpha_1| < |\alpha_2| \end{array} \tag{71}$$

$$\Delta\omega = |\omega_2 - \omega_1| . \tag{72}$$

The effective signal-to-noise ratio for the order determination problem is given by

$$SNR_{eff}^{(I)} = \frac{N}{\sigma^2} \min(|\alpha_1|^2, |\alpha_2|^2) \, \eta(\rho,\phi,\Delta\omega) = SNR_{amp} \, \eta(\rho,\phi,\Delta\omega) \tag{73}$$

where

$$\eta(\rho,\phi,\Delta\omega) = \frac{|\sqrt{1 + \rho^2 + 2\rho\cos(\phi + \gamma\Delta\omega)} - \sqrt{1 + \rho^2 + 2\rho\cos(\phi - \gamma\Delta\omega)}|^2}{4\rho^2} . \tag{74}$$

By comparing (73) with (64) we see that the effective signal-to-noise ratio

for closely spaced frequencies is reduced from the amplitude signal-to-noise ratio SNR_{amp} by the factor $\eta(\rho, \phi, \Delta\omega)$. Thus, in order to maintain a given effective signal-to-noise ratio for order determination, the amplitude signal-to-noise ratio must be increased by the reciprocal of the reduction factor (as compared to the case of widely separated frequencies).

The effective signal-to-noise ratio reduction factor $\eta(\rho, \phi, \Delta\omega)$ is plotted in figure E-1 versus $\Delta\omega$ for various values of ρ and ϕ.* The general behavior of this function may also be examined analytically by considering a second-order Taylor expansion of (74) for small $\Delta\omega$:

$$\eta(\rho, \phi, \Delta\omega) \approx \frac{(\gamma\Delta\omega)^2 \sin^2 \phi}{1 + \rho^2 + 2\rho\cos\phi} . \tag{75}$$

(iv) Closely Spaced Frequencies: Second-Order Signal with

0° or 180° Relative Phase

According to (70), if the window spectrum is given by the approximation (65), the minimum first-order approximation distance is exactly zero for a second-order signal with relative phase $\phi = 0$ or π; i.e., the second-order signal is indistinguishable from the best first-order model of it. When the actual window spectrum (46) is used in place of the approximation, there is a small difference between the two models.

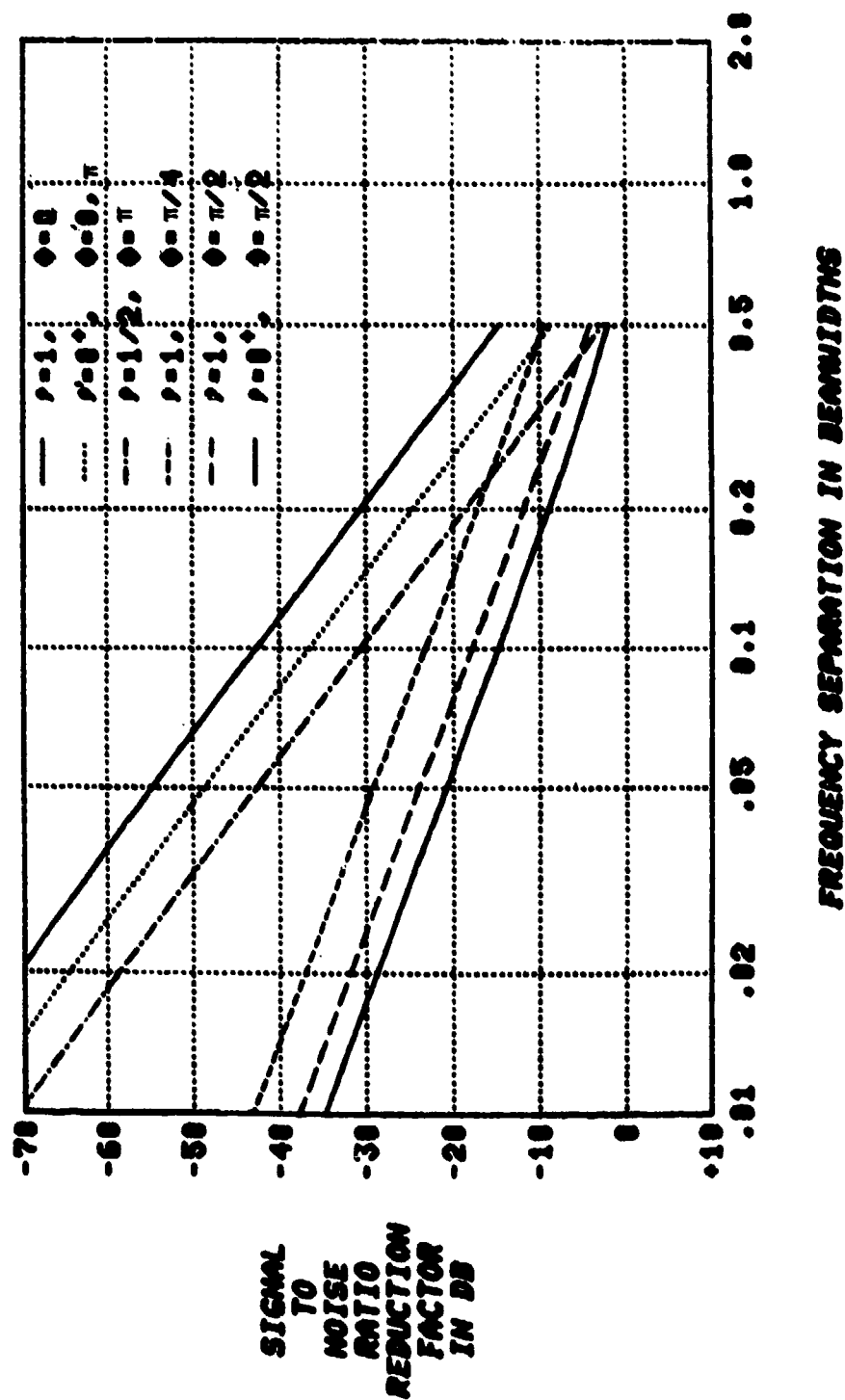*The curves for $\phi = 0$, π are calculated from (81), (82) of the next section rather than from (74).

Fig. E-1. Signal-to-noise ratio reduction factor, $10 \log_{10} \eta(\rho, \phi, \Delta\omega)$, as a function of frequency separation $\Delta f = \Delta\omega/2\pi$.

To calculate the actual minimum first-order approximation distance we first add on a fourth-order correction to the window spectrum approximation,

$$W(\omega) \approx \cos\gamma\omega + \gamma' \omega^4 \tag{76}$$

where

$$\gamma' = \frac{1}{4320} \left(1 - \frac{1}{N^2}\right) \left(1 - \frac{4}{N^2}\right) . \tag{77}$$

Then we recalculate the matrix in (67) as

$$(W_{II} - W_{LI}^+ W_{LL}^{-1} W_{LI})_{ij} = \sin|\gamma(\tilde{\omega}_1 - \omega_i)| \; \sin|\gamma(\tilde{\omega}_1 - \omega_j)|$$

$$- \; |4(\tilde{\omega}_1 - \omega_i)^2 - 6(\tilde{\omega}_1 - \omega_i)(\tilde{\omega}_1 - \omega_j) + 4(\tilde{\omega}_1 - \omega_j)^2|$$

$$\times \; \gamma'(\tilde{\omega}_1 - \omega_i)(\tilde{\omega}_1 - \omega_j)$$

$$+ \text{ higher order terms.} \tag{78}$$

It is assumed that the minimizing approximation frequency $\tilde{\omega}_1$ is chosen to minimize the contribution to $\tilde{\delta}^{(1)}$ of the dominant first term in (78); i.e., $\tilde{\omega}_1$ is selected according to the criterion following (68). For real α_1, α_2 and small separation frequency $\Delta\omega$, the optimum $\tilde{\omega}_1$ can be expressed as

$$\tilde{\omega}_1 \approx \frac{\alpha_1\omega_1 + \alpha_2\omega_2}{\alpha_1 + \alpha_2} . \tag{79}$$

Therefore, the minimum first-order approximation distance is evaluated as

$$|\tilde{\delta}_{min}^{(1)}|^2 = \underline{\alpha}^+(W_{II} - W_{LI}^+ W_{LL}^{-1} W_{LI})\underline{\alpha}$$

$$= \begin{cases} \min(\alpha_1^2, \alpha_2^2) \dfrac{6\gamma'(\Delta\omega)^4}{(1+\rho)^2}, & \phi = 0 \\[4mm] \min(\alpha_1^2, \alpha_2^2) \dfrac{6\gamma'(\Delta\omega)^4}{(1-\rho)^2}, & \phi = \pi \end{cases} \qquad (80)$$

and the effective signal-to-noise ratio reduction factor is given by

$$\eta(\rho, 0, \Delta\omega) = \frac{6\gamma'(\Delta\omega)^4}{(1+\rho)^2} \qquad (81)$$

$$\eta(\rho, \pi, \Delta\omega) = \frac{6\gamma'(\Delta\omega)^4}{(1-\rho)^2} . \qquad (82)$$

c. General Comments

An interesting feature of the analysis presented in this section is that none of it depends on an assumption of large sample size N. In fact, the window spectrum approximation (65) used for the closely spaced frequency case is exact when N = 2. This simple case can provide some insight into the mysterious $\sin^2\phi$ dependence of the effective signal-to-noise ratio for closely spaced frequencies.

We consider N = 2 and I = 2, and we seek to perform a measurement to decide whether the signal $\underline{s}$ is first order or second order. Instead of trying to measure the minimum approximation distances $\tilde{\delta}_{min}^{(L)}$, L = 1, 2, we use a simpler statistic based on the ratio s_2/s_1 of the two signal samples. If the signal were first order we would find that $|s_2/s_1| = 1$ exactly. For a second

-order signal,

$$\left| 1 - |s_2/s_1|^2 \right| = \left| 1 - \left| \frac{\alpha_1 \beta_1^{1/2} + \alpha_2 \beta_2^{1/2}}{\alpha_1 \beta_1^{-1/2} + \alpha_2 \beta_2^{-1/2}} \right|^2 \right|$$

$$= \left| 1 - \frac{|\alpha_1|^2 + |\alpha_2|^2 + 2\mathrm{Re}[\alpha_1^* \alpha_2 e^{j(\omega_2 - \omega_1)/2}]}{|\alpha_1|^2 + |\alpha_2|^2 + 2\mathrm{Re}[\alpha_1^* \alpha_2 e^{-j(\omega_2 - \omega_1)/2}]} \right|$$

$$= \frac{4\,\mathrm{Im}(\alpha_1^* \alpha_2)\,\sin(\omega_2 - \omega_1)/2}{|\alpha_1|^2 + |\alpha_2|^2 + 2\mathrm{Re}[\alpha_1^* \alpha_2 e^{-j(\omega_2 - \omega_1)/2}]}$$

$$\approx \frac{2\rho\,\Delta\omega\,|\sin\phi|}{1 + \rho^2 + 2\rho\cos\phi} \quad \text{if } \Delta\omega \text{ is small .} \tag{83}$$

This is the same type of frequency and phase dependence exhibited by the minimum first-order approximation distance $\tilde{\delta}_{\min}^{(1)}$. In particular, if the relative phase is $0°$ or $180°$ it is impossible to distinguish the second-order signal from a first-order signal on the basis of amplitude variation.

It is significant, though, that this distinction is possible for other phases, because such resolution does not appear to exist under a general ARMA model. In our vector space formulation, a general ARMA model with poles not restricted to the unit circle corresponds to using a larger set B' of admissible representation vectors, with

$$B' = \left\{ \underline{\beta} = [\beta_1, \ldots, \beta_N]^T : \beta_k = \beta^{k - \frac{N+1}{2}} \text{ for some complex } \beta, \right.$$

$$\text{not necessarily satisfying } |\beta| = 1\} \ . \qquad (84)$$

The minimum approximation distance $\hat{\delta}_{min}^{(L)}$ achieved by projections into L-dimensional subspaces spanned by vectors drawn from B' cannot be any greater than the $\hat{\delta}_{min}^{(L)}$ computed for B, and in fact it can be considerably smaller. A simple manifestation of this phenomenon is the failure under a general ARMA model of the amplitude comparison scheme described above. With admissible representation vectors drawn from B' instead of from B, it is not necessarily true that $|s_2/s_1| = 1$ for a first-order signal and thus the basis for deciding between first order and second order is lost.

6. THE MAXIMUM LIKELIHOOD ESTIMATOR FOR THE MINIMUM LTH ORDER APPROXIMATION DISTANCE

In Section 3 we calculated the Cramer-Rao accuracy bound on any unbiased estimate of the minimum L^{th}-order approximation distance, when the signal is observed in white Gaussian noise as in (1) - (13). In Section 4 we related the accuracy of the approximation distance estimate to the error probability achieved by a signal order estimate derived from the estimates of the minimum approximation distances. In this section we derive the maximum likelihood estimator for the minimum approximation distances and examine the error probability achieved by the corresponding estimate of the signal order.

Since the observed data $\underline{x}$ consists of the signal $\underline{s}$ plus white Gaussian noise $\underline{n}$, the maximum likelihood estimate $\hat{\underline{s}}$ of the signal based on $\underline{x}$ is calculated from the combination of unknown signal parameters which

minimizes $|\underline{x} - \underline{s}|$. For an Lth-order model, the maximum likelihood estimate is obviously

$$\hat{\underline{s}} = P_L(\hat{\underline{\beta}}_1, \ldots, \hat{\underline{\beta}}_L) \underline{x} \tag{85}$$

where $\hat{\underline{\beta}}_1, \ldots, \hat{\underline{\beta}}_L, \hat{\underline{\beta}}_i \in B$, are admissible representation vectors which minimize $|\underline{x} - P_L(\underline{\beta}_1, \ldots, \underline{\beta}_L)\underline{x}|$. The resulting minimum data approximation distance is the maximum likelihood estimate of the minimum L^{th}-order signal approximation distance $\tilde{\delta}_{min}^{(L)}$, i.e.,

$$\hat{\delta}_{ML}^{(L)} = \min_{\underline{\beta}_i \in B} |\underline{x} - P_L(\underline{\beta}_1, \ldots, \underline{\beta}_L)\underline{x}| = |\underline{x} - P_L(\hat{\underline{\beta}}_1, \ldots, \hat{\underline{\beta}}_L)\underline{x}| \ . \tag{86}$$

The maximum likelihood approximation distance estimator $\hat{\delta}_{ML}^{(L)}$ is a monotonically decreasing function of the approximation order L (as desired) for any value of the observed data vector $\underline{x}$. It is difficult to determine the exact statistical behavior of $\hat{\delta}_{ML}^{(L)}$ because of the dependence of the minimizing projection in (86) on the observed data.

It is possible to obtain analytically useful bounds on the false alarm and miss probabilities $P_F^{(I)}$, $P_M^{(I)}$ attained by the order determination algorithm of Section 4 applied to the maximum likelihood estimator $\hat{\delta}_{ML}^{(L)}$. For any $\underline{x}$, the minimum data approximation distance $\hat{\delta}_{ML}^{(L)}$ in (86) is obviously upper bounded by the approximation distance $\hat{\delta}_o^{(L)}$ that would result from using the L representation vectors $\tilde{\underline{\beta}}_1^o, \ldots, \tilde{\underline{\beta}}_L^o$ which best approximate the true signal (uncorrupted by noise), i.e.,

$$|\hat{\delta}_{ML}^{(L)}|^2 < |\hat{\delta}_o^{(L)}|^2 \equiv |\underline{x} - P_L(\tilde{\underline{\beta}}_1^o, \ldots, \tilde{\underline{\beta}}_L^o)\underline{x}|^2$$

$$= |\delta_{min}^{(L)} \underline{s}^{\perp}(\underline{\tilde{\beta}}_1^{o}, \ldots, \underline{\tilde{\beta}}_L^{o}) + \underline{n}^{\perp}(\underline{\tilde{\beta}}_1^{o}, \ldots, \underline{\tilde{\beta}}_L^{o})|^2 \qquad (87)$$

where $\underline{n}^{\perp}(\underline{\tilde{\beta}}_1^{o}, \ldots, \underline{\tilde{\beta}}_L^{o})$ is the orthogonal projection of the noise vector,

$$\underline{n}^{\perp}(\underline{\tilde{\beta}}_1^{o}, \ldots, \underline{\tilde{\beta}}_L^{o}) = \underline{n} - P_L(\underline{\tilde{\beta}}_1^{o}, \ldots, \underline{\tilde{\beta}}_L^{o})\underline{n} \qquad (88)$$

and $\underline{s}^{\perp}(\underline{\beta}_1^{o}, \ldots, \underline{\beta}_L^{o})$ is a unit vector in the direction of the orthogonal projection of the signal vector,*

$$\underline{s}^{\perp}(\underline{\tilde{\beta}}_1^{o}, \ldots, \underline{\tilde{\beta}}_L^{o}) = \lfloor \underline{s} - P_L(\underline{\tilde{\beta}}_1^{o}, \ldots, \underline{\tilde{\beta}}_L^{o}) \underline{s}\rfloor/\tilde{\delta}_{min}^{(L)} \quad . \qquad (89)$$

Thus, the upper bound on the square of the maximum likelihood approximation distance estimator is simply a noncentral (unless $\tilde{\delta}_{min}^{(L)} = 0$) chi-squared random variable with $2(N-L)$ degrees of freedom.

The relevant approximation distance estimate for computing the false alarm probability $P_F^{(I)}$ is the Ith order estimate $\hat{\delta}_{ML}^{(I)}$. Since the true minimum I^{th}-order approximation distance is obviously $\tilde{\delta}_{min}^{(I)} = 0$ (I = true signal order), the upper bound (87) on the square of the Ith-order maximum likelihood estimator is simply the (central) chi-squared random variable $|\underline{n}^{\perp}(\underline{\tilde{\beta}}_1^{o}, \ldots, \underline{\tilde{\beta}}_L^{o})|^2$. Therefore, for any order determination threshold ϵ, the false alarm probability is upper bounded by

$$P_F^{(I)} = Pr|\hat{\delta}_{ML}^{(I)} > \epsilon| < Pr||\underline{n}^{\perp}(\underline{\tilde{\beta}}_1^{o}, \ldots, \underline{\tilde{\beta}}_L^{o})|^2 > \epsilon^2| \quad . \qquad (90)$$

As a function of ϵ^2 the upper bound is a chi-squared probability distribu-

tion function. Evaluation leads to

$$P_F^{(I)} < e^{-N\epsilon^2/\sigma^2} \sum_{m=0}^{N-I-1} \frac{1}{m!} (N\epsilon^2/\sigma^2)^m$$

$$< e^{-N\epsilon^2/\sigma^2} \sum_{m=0}^{N-1} \frac{1}{m!} (N\epsilon^2/\sigma^2)^m \quad . \tag{91}$$

The upper bound in (91) is plotted in figure E-2 as a function of the "threshold-to-noise" ratio $N\epsilon^2/\sigma^2$, for all possible sample sizes from $N = 2$ to $N = 32$. It is seen that the false alarm probability goes to zero exponentially as the threshold-to-noise ratio is _increased_. Since the (second) bound in (91) is independent of I, an appropriate order determination threshold to insure a tolerable false alarm probability can be selected from Fig. E-2 without any prior knowledge of the true signal order.

Analysis of the miss probability $P_M^{(I)}$ is more difficult. To begin with, the upper bound (87) on the maximum likelihood estimator of the minimum approximation distance leads to a _lower_ bound on the miss probability, i.e.,

$$P_M^{(I)} = Pr\{\hat{\delta}_{ML}^{(I-1)} < \epsilon\}$$

$$\tag{92}$$

$$> Pr\{|\tilde{\delta}_{min}^{(I-1)} \underline{s}^\perp(\tilde{\underline{\beta}}_1^o, \ldots, \tilde{\underline{\beta}}_{I-1}^o) + \underline{n}^\perp(\tilde{\underline{\beta}}_1^o, \ldots, \tilde{\underline{\beta}}_{I-1}^o)|^2 < \epsilon^2\}.$$

If the signal-to-noise ratio is high enough to produce good maximum likelihood estimates $\hat{\underline{\beta}}_1, \ldots, \hat{\underline{\beta}}_{I-1}$ of the optimum $(I-1)^{th}$-order approximation vectors $\tilde{\underline{\beta}}_1^o, \ldots, \tilde{\underline{\beta}}_{I-1}^o$, then the lower bound in (92) will provide a good approximation of the actual miss probability.
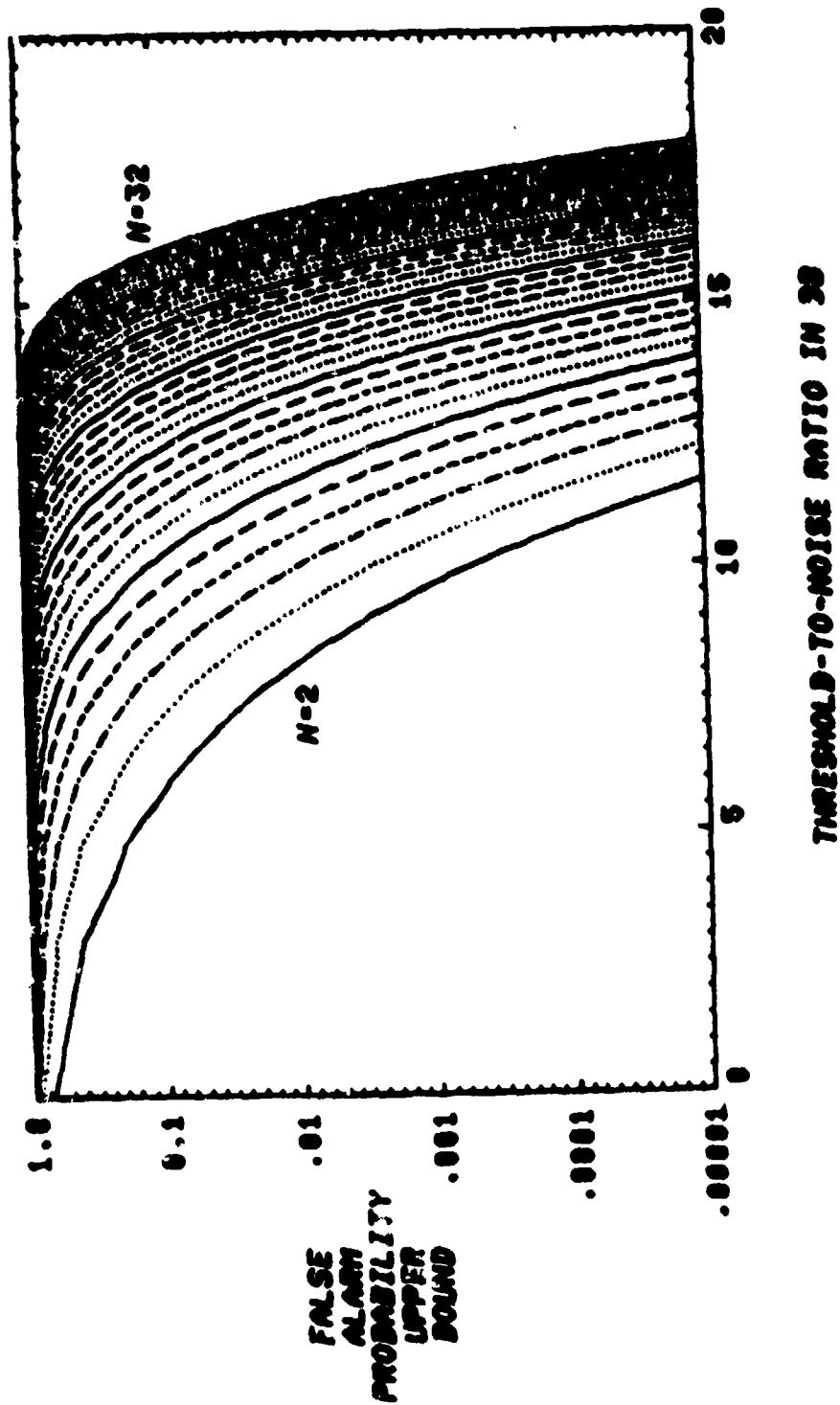
Fig. E-2. Upper bound on false alarm probability $P_F^{(I)}$ as a function of the decision threshold-to-noise ratio $(10 \log_{10} N\epsilon^2/\sigma^2)$.

As a function of ϵ^2 the lower bound is a noncentral chi-squared distribution function. Evaluation leads to an expansion of the form

$$P_M^{(I)} > \sum_{k=0}^{\infty} \frac{1}{k!} |SNR_{eff}^{(I)}|^k e^{-SNR_{eff}^{(I)}} \sum_{m=N-1+k+1}^{\infty} \frac{1}{m!} (N\epsilon^2/\sigma^2)^m e^{-N\epsilon^2/\sigma^2}$$

$$> \sum_{k=0}^{\infty} \frac{1}{k!} |SNR_{eff}^{(I)}|^k e^{-SNR_{eff}^{(I)}} \sum_{m=N+k+1}^{\infty} \frac{1}{m!} (N\epsilon^2/\sigma^2)^m e^{-N\epsilon^2/\sigma^2}. \tag{93}$$

Both summations in (93) involve Poisson probabilities. The result of the inner summation goes to zero rapidly as the summation limit (N+k+1) is increased beyond the corresponding Poisson mean, $N\epsilon^2/\sigma^2$. Such large values of k are overwhelmingly probable whenever the Poisson mean in the outer summation, $SNR_{eff}^{(I)}$, is much greater than $N\epsilon^2/\sigma^2$. Thus, a small miss probability may be expected whenever the effective signal-to-noise ratio greatly exceeds the threshold-to-noise ratio.

Figures E-3a, E-3b, and E-3c plot the lower bound on the miss probability versus the threshold-to-noise ratio for all sample sizes from N=2 to N=32 and three values of the effective signal-to-noise ratio, $SNR_{eff}^{(I)}$ = 5, 10, 20. It is seen that the miss probability bound decreases exponentially as the threshold-to-noise ratio is _decreased_.

The results of figures E-2 and E-3 can be combined to determine the trade-off between false alarm probability and miss probability as the decision threshold is varied. Figures E-4a, E-4b, and E-4c plot the lower bound (93) on the miss probability versus the upper bound (91) on the false alarm probability for N = 2, ..., 32 and $SNR_{eff}^{(I)}$ = 5, 10, 20. It is seen that the error probability trade-off curves for fixed effective signal-to-noise ratio
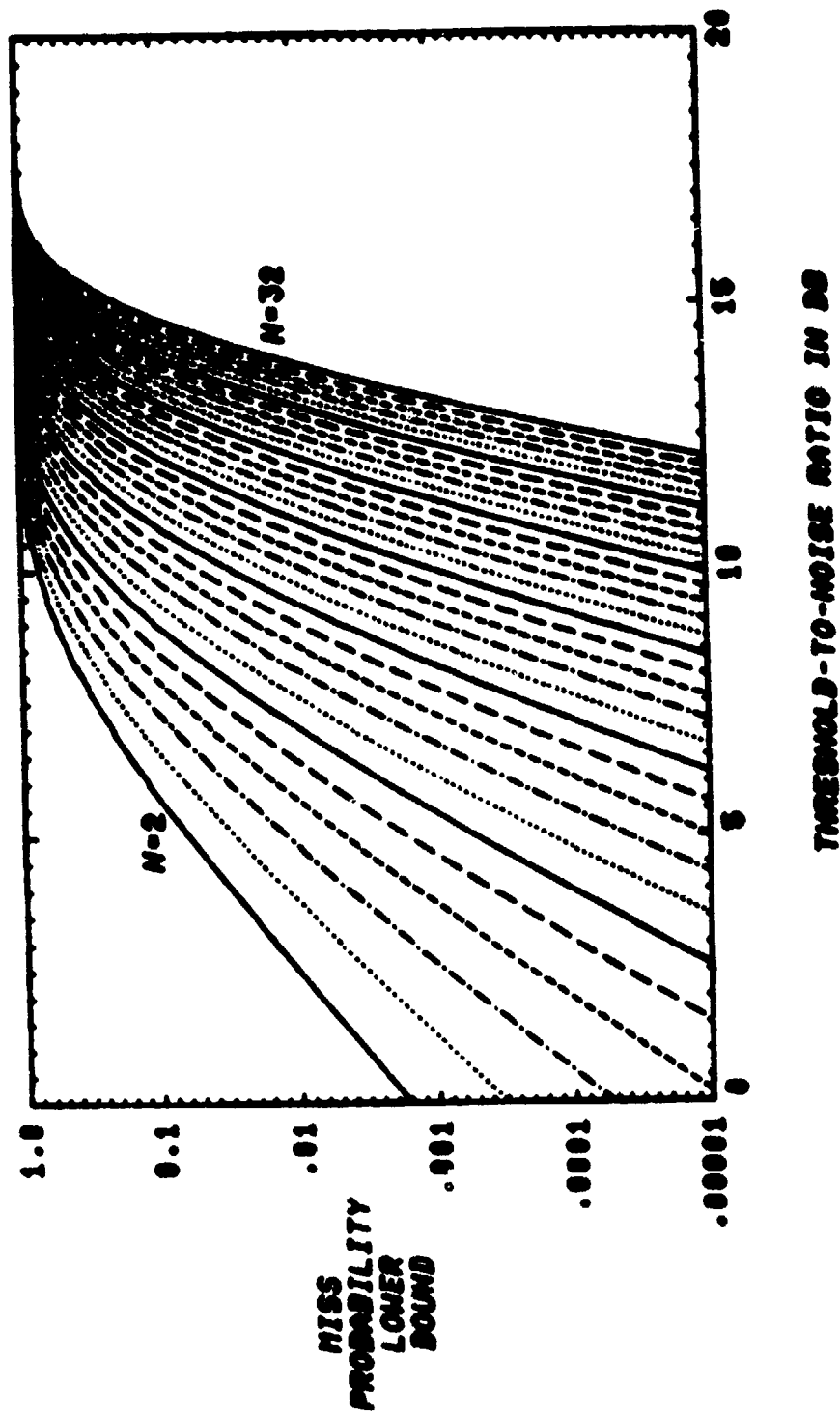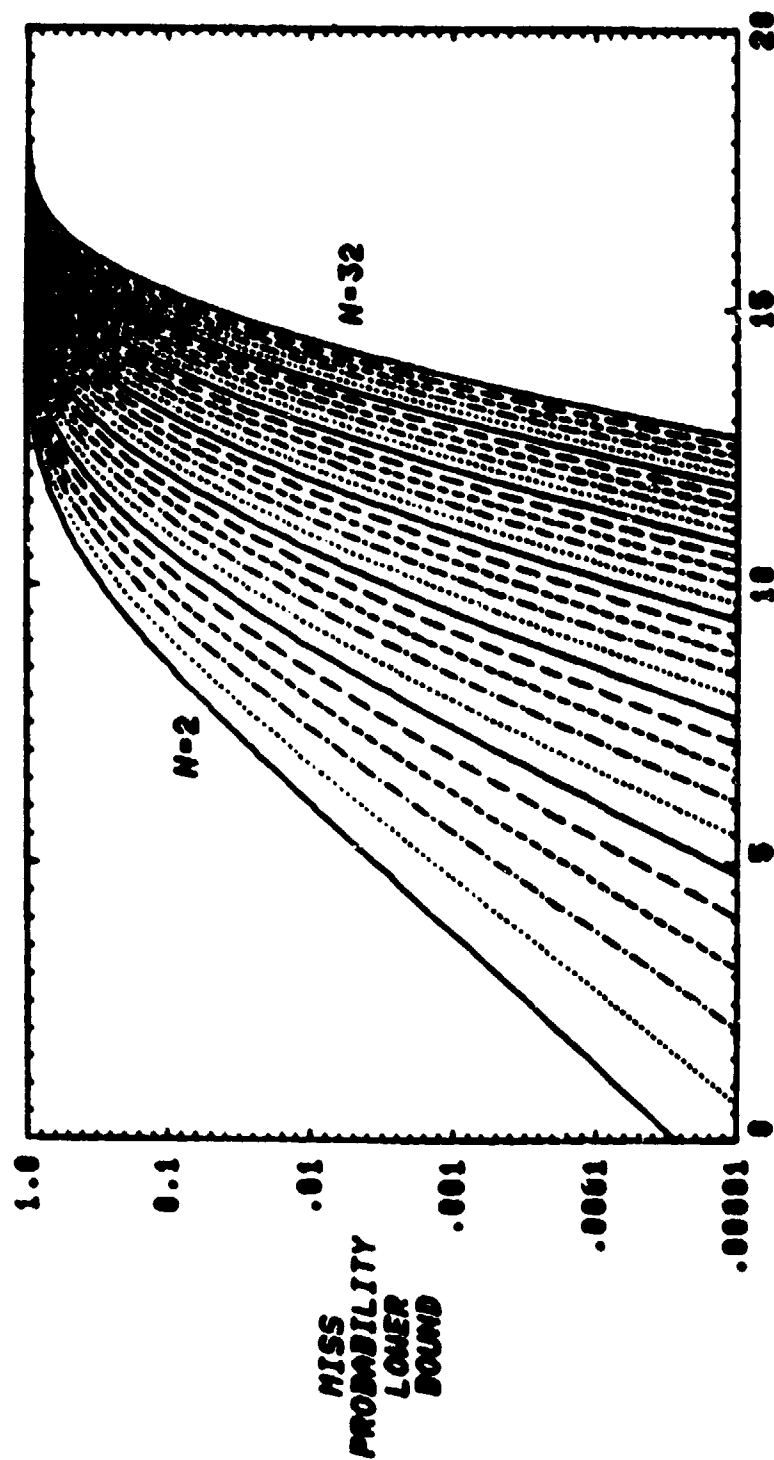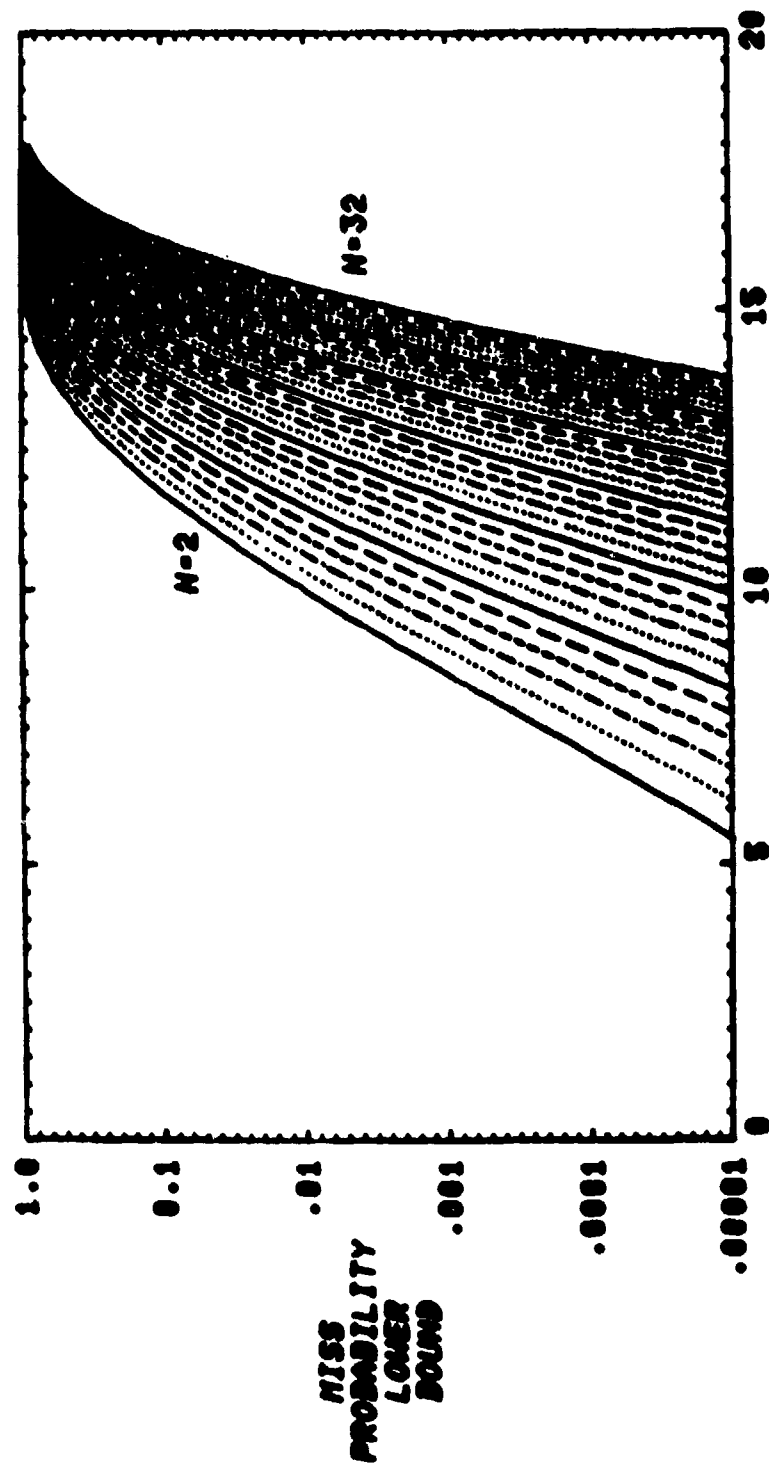
Fig. E-3a. Lower bound on miss probability $P_M^{(I)}$ as a function of the decision threshold-to-noise ratio ($10 \log_{10} N\epsilon^2/\sigma^2$) for an effective signal-to-noise ratio of $SNR_{eff}^{(I)} = 5$ (7 dB).

Fig. E-3b. Lower bound on miss probability $P_M^{(I)}$ as a function of the decision threshold-to-noise ratio ($10 \log_{10} N \varepsilon^2/\sigma^2$) for an effective signal-to-noise ratio of $SNR_{eff}^{(I)} = 10$ (10 dB).

Fig. E-3c. Lower bound on miss probability $P_M^{(I)}$ as a function of the decision threshold-to-noise ratio ($10 \log_{10} N\epsilon^2/\sigma^2$) for an effective signal-to-noise ratio of $SNR_{eff}^{(I)} = 20$ (13 dB).
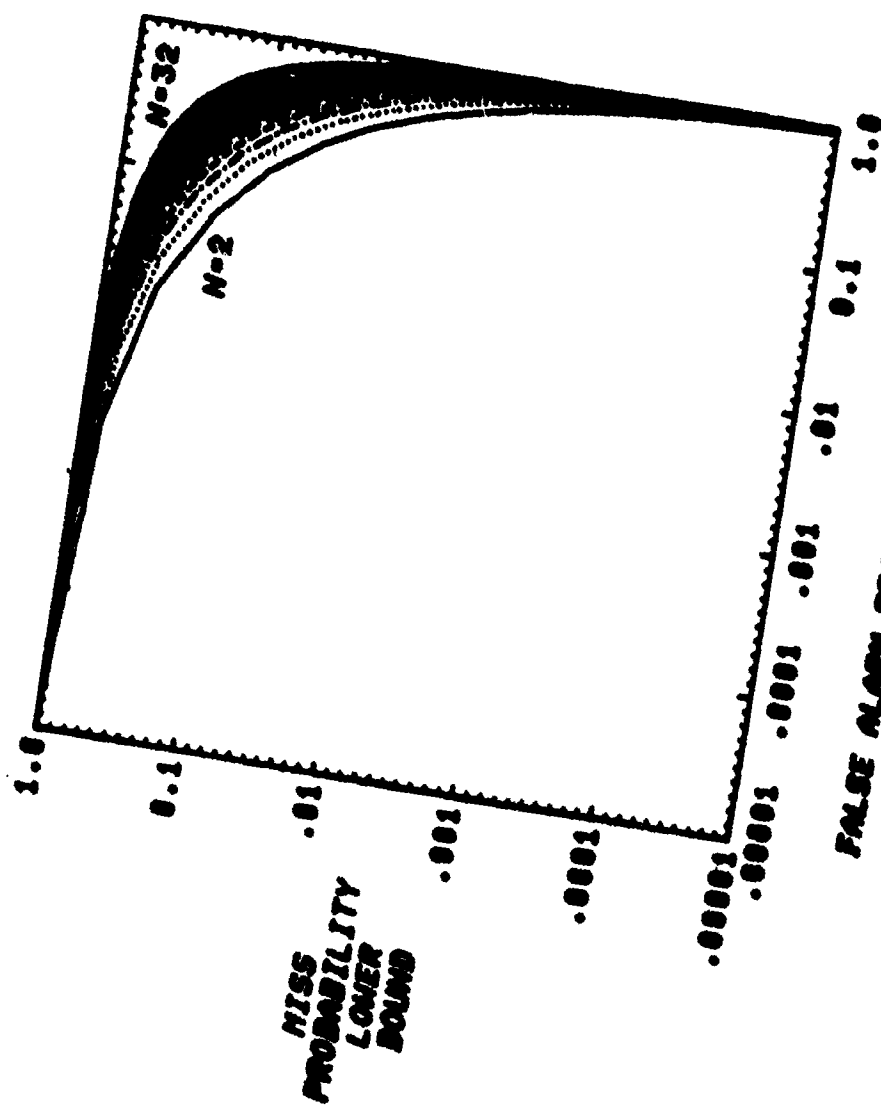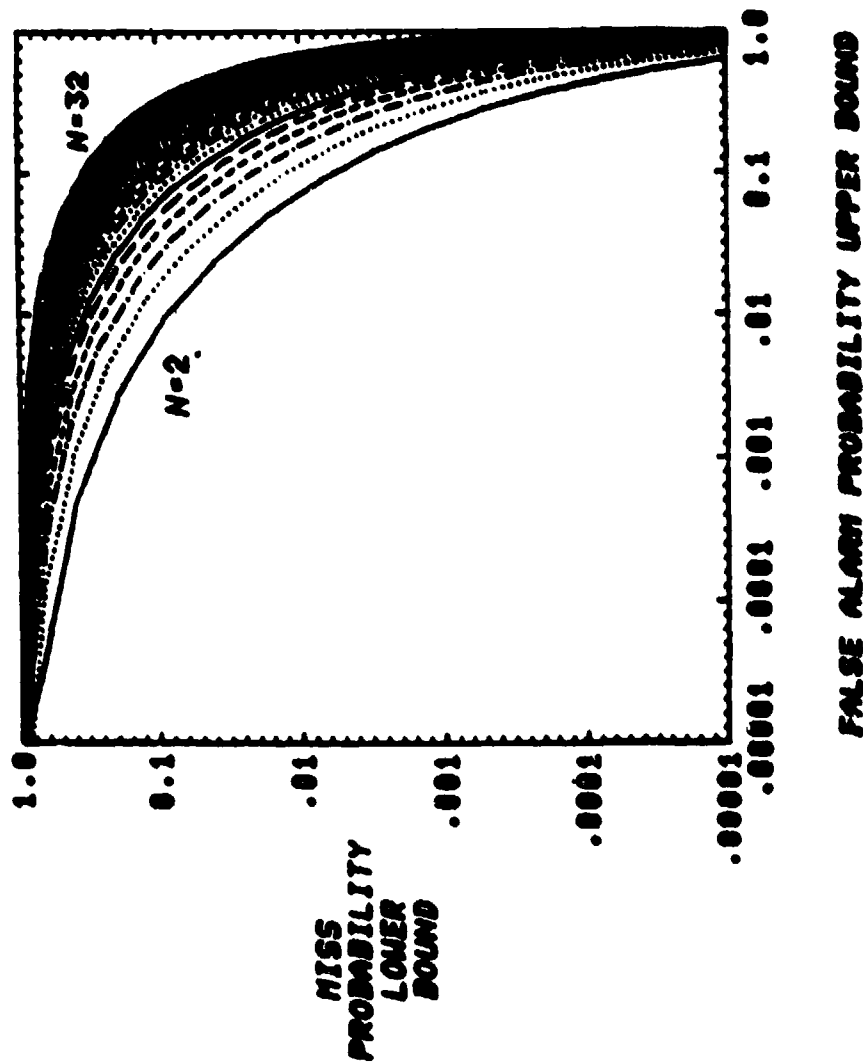
Fig. E-4a. Tradeoff relation between the lower bound on the miss probability $P_M^{(I)}$ and the upper bound on the false alarm probability $P_F^{(I)}$ for an effective signal-to-noise ratio of $SNR_{eff}^{(I)} = 5$ (7 dB).

E-35

FALSE ALARM PROBABILITY UPPER BOUND

Fig. E-4b. Tradeoff relation between the lower bound on the miss probability $P_M^{(I)}$ and the upper bound on the false alarm probability $P_F^{(I)}$ for an effective signal-to-noise ratio of $SNR_{eff}^{(I)} = 10$ (10 dB).
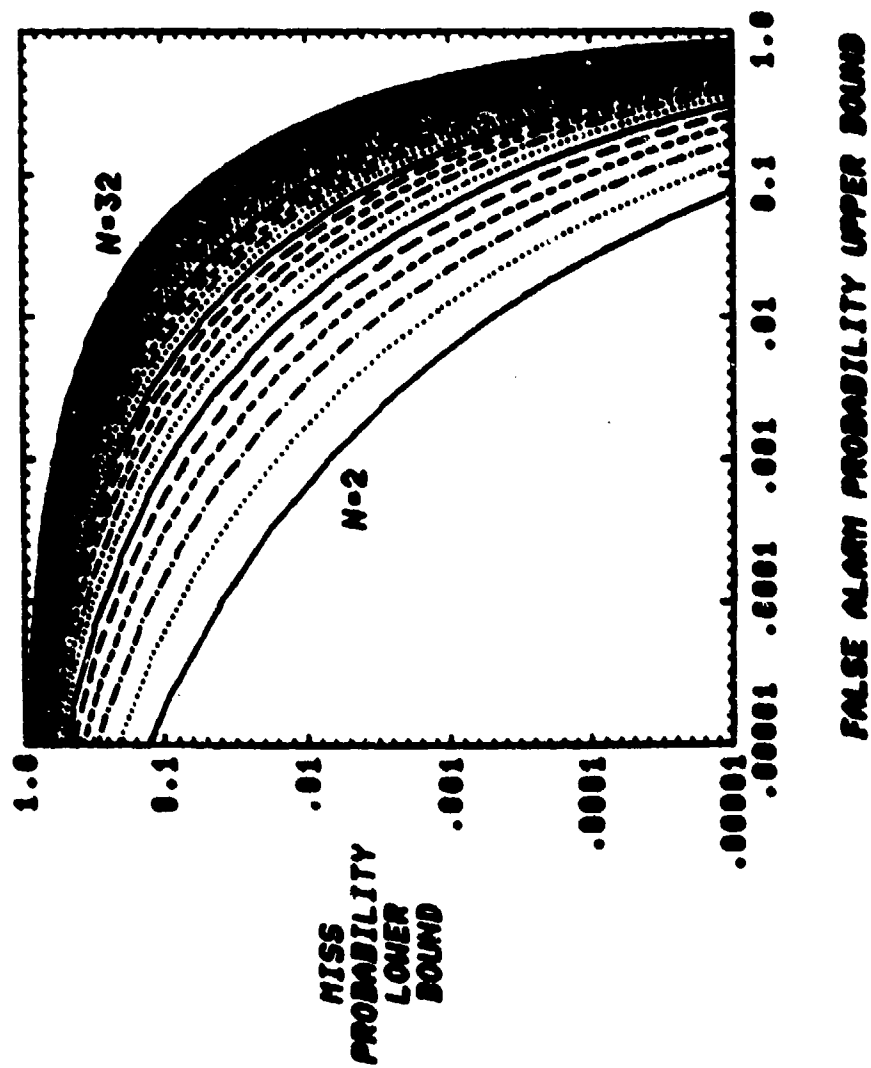
FALSE ALARM PROBABILITY UPPER BOUND

Fig. E-4c. Tradeoff relation between the lower bound on the miss probability $P_M^{(I)}$ and the upper bound on the false alarm probability $P_F^{(I)}$ for an effective signal-to-noise ratio of $SNR_{eff}^{(I)} = 20$ (13 dB).

are not highly dependent on the sample size N, and in fact they approach a limiting trade-off function as $N \to \infty$.

If the errors of overestimating and underestimating the true signal order are equally undesirable, then the threshold-to-noise ratio should be chosen to minimize the total error probability (i.e., the sum of the false alarm and miss probabilities). Figure E-5 plots the approximate optimum error probability, obtained by minimizing the sum of the false alarm and miss probability bounds, as a function of the effective signal-to-noise ratio $SNR_{eff}^{(2)}$. It is seen that effective signal-to-noise ratios between 10 dB and 20 dB are sufficient to produce error probabilities between 10^{-1} and 10^{-5}.

7. COMPUTER SIMULATION OF THE PERFORMANCE OF THE ORDER ESTIMATOR

The order estimation algorithm described in Sections 4 and 6 was simulated with synthetic data. One hundred independent noise realizations $\underline{n}_q$, $q = 1, \ldots, 100$, were generated, each with a sample size $N = 9$. Next, a known second-order signal $\underline{s}$ was added to each of the noise realizations, resulting in 100 data sets $\underline{x}_q$, $q = 1, \ldots, 100$. Maximum likelihood estimators $\hat{\delta}_q^{(L)}$, $L = 1, 2$ were derived from each data set $\underline{x}_q$ according to the formula (86). This procedure was repeated for signals with various amplitudes, phases, and frequencies. The same noise data was used for each signal $\underline{s}$.

A decision threshold ε was selected to guarantee a false alarm probability less than 0.05; according to figure E-2 the required threshold-to-noise ratio is $N\varepsilon^2/\sigma^2 = 14$ (i.e., 11.5 dB). Each second-order signal $\underline{s}$ is considered to have been resolved if $\hat{\delta}_q^{(2)} < \varepsilon < \hat{\delta}_q^{(1)}$ for at least 90% of the data sets $\underline{x}_q$. Figure E-6 presents the signal-to-noise ratios and frequency

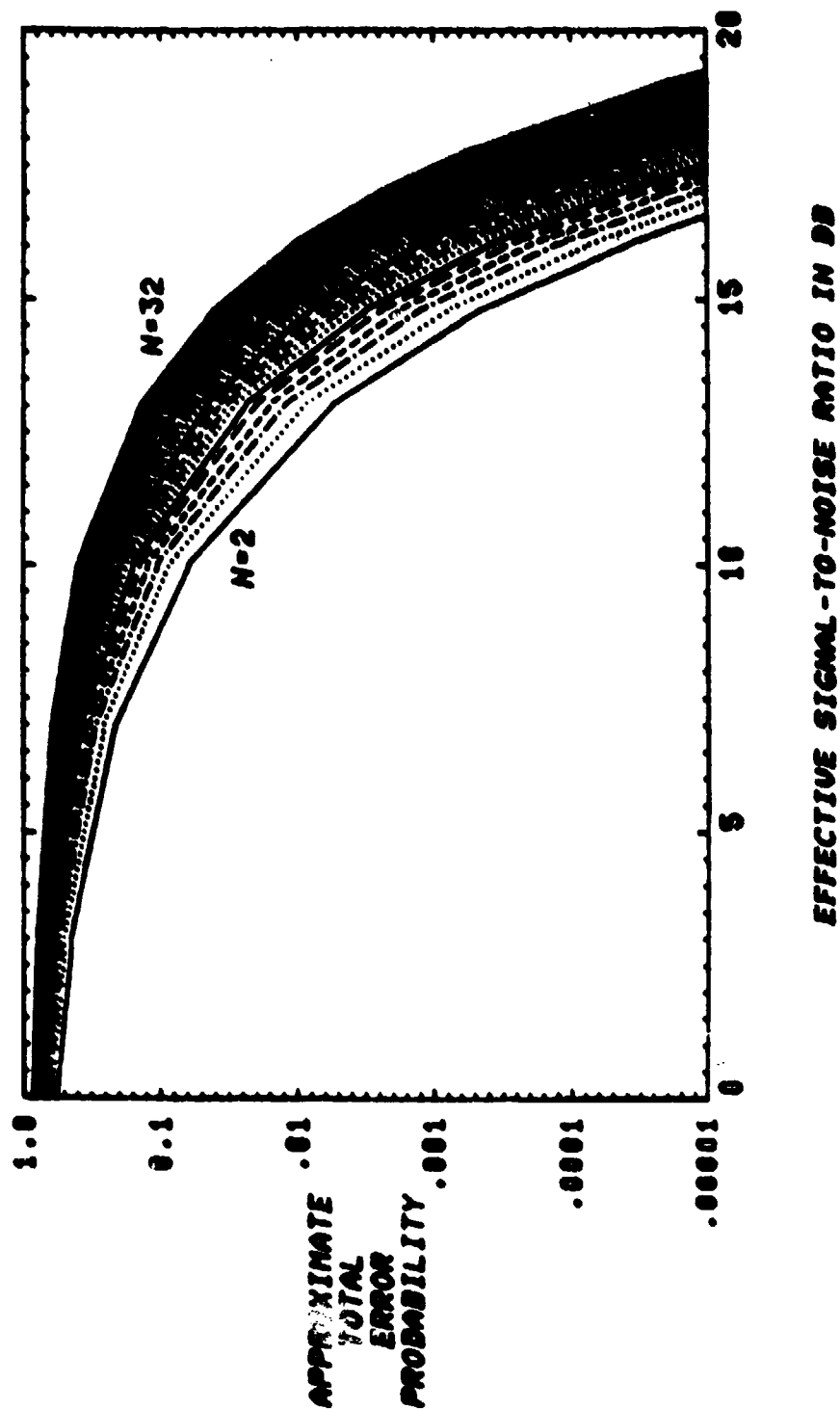**EFFECTIVE SIGNAL-TO-NOISE RATIO IN DB**

Fig. E-5. Approximate total error probability $P_F^{(I)} + P_M^{(I)}$ as a function of the effective signal-to-noise ratio $10 \log_{10} SNR_{eff}$.

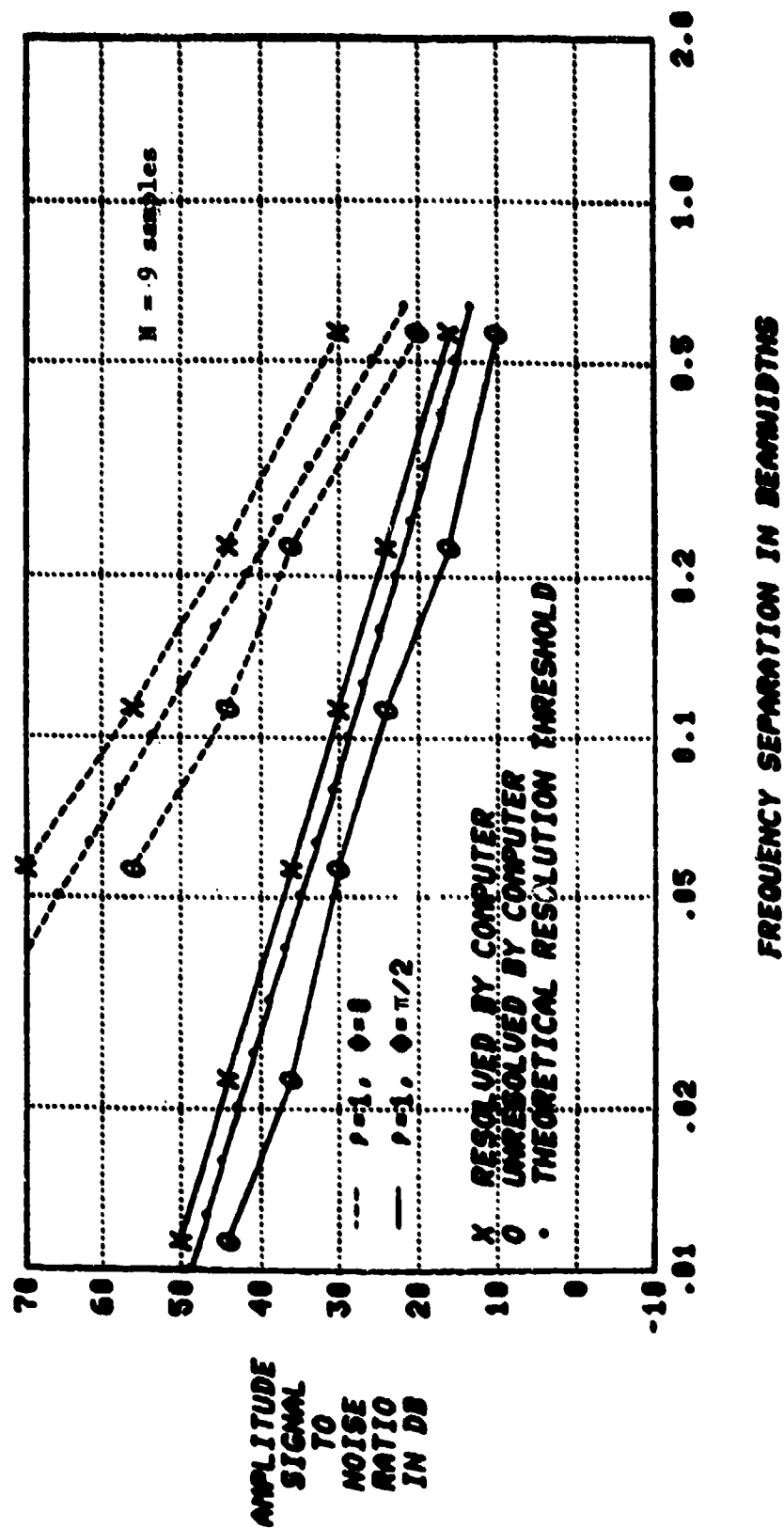**Fig. E-6.** Amplitude signal-to-noise ratio $10 \log_{10} SNR_{amp}$ required to resolve second-order signals with closely spaced frequencies ($\Delta f = \Delta\omega/2\pi$).

separations of the signals that were resolved and those that were not resolved. In the figure the separation coordinate is $\Delta f = \frac{\Delta \omega}{2\pi}$, and the plotted signal-to-noise ratio is $\mathrm{SNR}_{amp} = N \min(|\alpha_1|^2, |\alpha_2|^2)/\sigma^2$. Results are presented for two cases, $\phi = 0, \frac{\pi}{2}$, with $\rho = 1$. Three curves are shown for each phase value. The upper curve connects points (denoted by 'X') for which the order determination was successful, and the lower curve connects points (denoted by 'O') for which the true order was not estimated. The middle curve is the value of SNR_{amp} required to yield an <u>effective</u> signal-to-noise ratio $\mathrm{SNR}_{eff}^{(2)} = 12.5$ (i.e., 11 dB). According to figure E-5, this is the effective signal-to-noise ratio required to achieve an approximate total error probability of 0.1.

APPENDIX F

LOSS OF RESOLUTION WITH THE "STANDARD" MLM
METHOD WITH COHERENT MULTIPATH

In this appendix, we provide an explanation for the loss of resolution with the "normal" MLM technique when coherent multipath is present. By "normal" technique, we mean the spectrum estimate

$$P_{MLM}(\theta) = (\underline{v}^H \hat{\underline{R}}^{-1} \underline{v})^{-1} \qquad (F-1)$$

of Eq. (4-6) where the covariance estimate $\hat{R}$ is time averaged, but **not** spatially smoothed. With this particular estimate, it has been observed that the angular spectrum estimates for two coherent signals typically shows a single peak located between the two arrival angles and an apparent noise floor which is much closer to the spectrum peak levels than is the case with the actual noise level (see Fig. 4-9).

The explanation we give here is an adaptive nulling array interpretation due to Warren White[*] [13]. First we recall that one interpretation of the MLM method is that one seeks to minimize the output power from the array subject to the constraint that the gain in the steering direction, θ, is held fixed at unity gain. With uncorrelated signals, this is achieved by steering nulls towards the signals not at θ as was illustrated in Fig. 4-8.

However, with coherent signals (e.g., multipath), the array output power is no longer minimized by steering nulls to the signal angles other than θ. To see this, consider the case of 2 signals at angles θ_1 and θ_2 with amplitudes S_1 and $S_2 = S_1 \rho e^{j\phi}$. If we wish to estimate the power at angle θ_1 and are given the ensemble covariance for uncorrelated signals (so we can null out the second signal), the output power will be

$$P_{out} \sim S_1^2 + P_N$$

---

[*]A more mathematical discussion is available in Cantoni and Godara [88].

where P_N = array output noise power ($\approx$ noise power at an element/no. of elements). However, in the coherent case, if we have a weight gain of $-\rho e^{j\phi}$ in the direction of θ_2 then the array output will be

$$P(\theta_1) = <|S_1 + S_2 (-\rho e^{j\phi}) + n |^2> \approx P_N .$$

Similarly, when we estimate the power at angle θ_2, with uncorrelated signals, a null is formed in the direction of θ_1 so that

$$P(\theta_2) \sim S_2^2 + P_N$$

whereas with the coherent signals a weight gain of $-(1/\rho) e^{j\phi}$ in the direction of θ_1 yields:

$$P(\theta_2) = < | S_2 + S_1(- \frac{1}{\rho} e^{j\phi}) + n|^2> \sim P_N .$$

With uncorrelated signals, the power estimates at the two arrival angles are biased upward by P_N whereas with coherent signals, the arrival-angle power estimates are close to P_N. Thus, with coherent signals and a high SNR, the MLM power estimates at angles θ_1 and θ_2 corresponding to the actual plane-wave arrival angles are significantly lower than they should be.

The resolution performance depends on the power estimates at angles θ, other than the arrival angles. With incoherent signals, nulls are placed at angles θ_1 and θ_2 (recall Fig. 4-8) so that the power estimate is

$$P(\theta) \approx P_N (1+\alpha) \qquad \theta \neq \theta_1 \text{ or } \theta_2.$$

The parameter α can be significantly greater than 0 due to the very high (e.g., +20 dB) "sidelobes" of the steered pattern at angles other than θ, θ_1 or θ_2 (Fig. 4-10). But, as long as S_1^2 and S_2^2 are $> P_N$, there will be sharp peaks in the MLM spectrum estimate at θ_1 and θ_2.

With coherent signals, it is not necessary to point nulls in the directions of θ_1 and θ_2 to eliminate the plane-wave signals. Rather, one only needs to have

$$F(\theta_1, \theta) = -\rho e^{j\phi} F(\theta_2, \theta)$$

where $F(\theta_1, \theta)$ is the array (voltage) gain at θ_1 when steering to angle θ. Thus, the array output power will be

$$P(\theta) \sim P_N$$

i.e., the estimated power is roughly constant at all angles.

When the plane-wave signals are cancelled, the output noise power depends on the modulus of the weighting vector. It turns out that the spectrum weight modulus when pointing at θ_1, θ_2 and between these two angles is larger than that which arises at angles well removed from θ_1 and θ_2. Thus, there is some structure to the resulting MLM spectrum estimates for coherent signals such that one could roughly infer the sector with plane wave arrivals. However, the spectrum estimate is significantly poorer in all respects to that afforded by conventional (Fourier) techniques.